AStA Advances in Statistical Analysis (2025) 109:529–555
https://doi.org/10.1007/s10182-025-00533-6

ORIGINAL PAPER

Basketball players performance measurement with algorithmic survival data analysis

Ambra Macis[1]

Received: 24 September 2023 / Accepted: 1 July 2025 / Published online: 28 August 2025

Abstract
Performance measurement is of paramount importance in the context of sports analytics. A great variety of data analysis methods has been exploited to this aim. All these proposals almost never include resorting to survival analysis techniques, although time-to-event data are suitable for addressing this issue. This work aims to identify the main achievements of a National Basketball Association player that affect the time it takes for him to exceed a given threshold of points. In order to identify nonlinear effects and possible interactions among the predictors, the analysis is carried out with machine learning methods, specifically survival trees and random survival forests.

Keywords Sport analytics · Performance · Survival trees · Random survival forests

1 Introduction

Sport analytics has become widespread in recent years for answering different questions concerning several fields. Among these, there is the performance analysis of players/teams, that is universally regarded as of utmost importance, as it affects several aspects related to the determination of game strategies, optimal training programs, scouting, and it is also related to other analytic aims such as, for example, the prediction of tournament/matches outcomes and the identification of factors that distinguish successful and unsuccessful teams (Zuccolotto and Manisera 2020).

From a methodological point of view, a wide variety of techniques has been employed for performance measurement, yet hardly ever those associated with survival data analysis techniques.

Indeed, they have been used for numerous aims, other than performance measurement. Just as an example, survival analysis is widely used in the injury setting

✉ Ambra Macis
ambra.macis@unibs.it

1 Department of Economics and Management, University of Brescia, Contrada S. Chiara, 50, 25122 Brescia, Italy

for identifying the risk factors for injury (Hopkins et al. 2007; Buist et al. 2010; Mahmood et al. 2014; Ekeland et al. 2020; Venturelli et al. 2011; Beynnon et al. 2005; Zumeta-Olaskoaga et al. 2021; Lu et al. 2022; Macis 2024) and recovery after injuries and sport-related concussions (Sochacki et al. 2019; Jack et al. 2019; Mai et al. 2017; Lawrence et al. 2018; Dekker et al. 2017; Nelson et al. 2016; Kontos et al. 2019; Howell et al. 2019). Moreover, survival analysis has been applied for determining the features that may influence the dropout of young athletes in many sports (Pion et al. 2015; Moulds et al. 2020; Smith and Weir 2022; Back et al. 2022) or for evaluating the career length of professional basketball players (Fynn and Sonnenschein 2012). Furthermore, it has been used for evaluating the effect of team performance in the dismissal of coaches (Wangrow et al. 2018; Tozetto et al. 2019). Then, other studies analyzed the duration of Olympic success (Csurilla and Fertő 2022; Gutiérrez et al. 2011) or whether Olympic medallists live longer than the general population (Clarke et al. 2012). Furthermore, survival analysis has been used for identifying the criterion that may affect the decision of a football coach of doing the first substitution during a match (Del Corral et al. 2008). Finally, other works analyzed the impact of performance indicators on the time when the first or the second goal is scored in football (Pratas et al. 2016; Nevo and Ritov 2013) or studied times between goals in ice hockey (Thomas 2007).

This work deals with basketball and aims to investigate which are the main achievements (e.g., attempted shots, double doubles and rebounds) of a National Basketball Association (NBA) player in the pre-All-Star Game (AS) season segment that affect the time it takes for him to exceed a given amount of points during the post-All-Star Game season segment. Until now, to the best of our knowledge, no studies have used survival analysis for studying athletes' performance. The only exception regards a previous study (Macis et al. 2023a, b) in which the authors used Cox regression and Lasso Cox for answering the above question. It emerged that having attempted more two-point shots in the previous season segment, having been selected for the All-Star Game, and having gained a higher number of double doubles, increase the probability of scoring more than a given amount of points (this threshold has been fixed ensuring a 50% of censoring in the sample) in a shorter time. Moreover, it resulted that steals were negatively associated with the survival outcome, suggesting that who is more involved in defense tends to gain less points. However, the other defense variables were not selected by the two models. Therefore, the role of defense should be further investigated (Macis et al. 2023a).

The problem under exam in this work, analyzed in the same framework of Macis et al. (2023a), is the presence of nonlinear relationships and interaction effects between the main NBA players' achievements and the probability of exceeding a given cutoff in a shorter time. This analysis was performed using two different machine learning algorithms, survival trees and random survival forests, two nonparametric methods that allow high flexibility taking into account nonlinear effects and possible interactions among the predictors without specifying them beforehand. Some preliminary analyses have been reported in Macis (2023).

The paper is organized as follows. In Sect. 2 the methodological details of the algorithms used in the application and of the metrics used for performance evaluation are shown; then, a description of the data and the study design are provided in Sect. 3. Following, the obtained results are shown. The paper ends with the final discussion.

2 Methods

Survival analysis aims to study the occurrence of an event of interest during a given period of time (follow-up). It is characterized by the presence of *censoring*. Many kinds of censoring can occur; in this study, right censoring is considered. It occurs when the event has not been observed for an observation during the follow-up, so that the only known thing is the last time he/she did not experience the event (Collett 2015). In this context, the i^{th} observation is defined by three elements:

- the event indicator δ_i that takes value 1 if the subject experienced the event and 0 otherwise;
- the observed time $\tau_i = \min(t_i, c_i)$, that is equal to the event time t_i if $\delta_i = 1$, or to the censoring time c_i if $\delta_i = 0$;
- the vector of observed covariates $\boldsymbol{x}_i$.

In this paper, machine learning methods for survival analysis have been used. Methodological details about survival trees and random survival forests are reported in Sects. 2.1 and 2.2, respectively.

2.1 Survival trees

Survival trees are an extension of classification and regression trees (CART) to censored data, with the aim of identifying, by a recursive partitioning of the covariate space, subgroups of subjects homogeneous in terms of survival probability. Several proposals have been advanced over the years (Gordon and Olshen 1985; Ciampi et al. 1986, 1987; Segal 1988; Davis and Anderson 1989; LeBlanc and Crowley 1992, 1993; Hothorn et al. 2006; Zeileis et al. 2008; Kundu and Ghosh 2021), but there is still not an algorithm recognized as the best one (Macis 2022). Two of the most used algorithms are Relative Risk Trees (RRTs) (LeBlanc and Crowley 1992) and Conditional Inference Trees (CITs) (Hothorn et al. 2006). One of the main advantages of these two algorithms is their flexibility and the interpretability of the obtained results. However, as a disadvantage, they are sensitive to noise in the data. Methodological details about RRTs and CITs have been provided in the following subsections.

2.1.1 Relative risk trees

Relative Risk Trees (LeBlanc and Crowley 1992) are based on the proportional hazards (PH) assumption and exploit the connection between the likelihoods of the proportional hazards and Poisson models for growing trees. To this extent, RRTs are equivalent to Poisson trees.

RRTs are grown using as splitting criterion the deviance, a within-node error measure that allows the use of the CART cost-complexity algorithm for pruning.

In detail, the proposed algorithm splits the covariate space into regions that maximize the reduction in one-step deviance realized by the split. Then, as in CART, the binary splitting procedure continues until a full-grown tree is obtained and there are only a few observations in each node. Finally, the tree is pruned using the cost-complexity algorithm (LeBlanc and Crowley 1992), which efficiently yields trees that perform best in terms of residual error (deviance) given their size:

$$R_\alpha(\mathscr{T}) = \sum_{k \in \tilde{\mathscr{T}}} R(k) + \alpha|\tilde{\mathscr{T}}| \,, \tag{1}$$

where $R(k)$ is the impurity of node k, α is a non-negative complexity parameter, and $\tilde{\mathscr{T}}$ is the set of terminal nodes of $\mathscr{T}$. The complexity parameter allows to control the trade-off between the goodness-of-fit of the model and its size. So, the optimally pruned trees are those that minimize the cost-complexity for given α's, and the smallest optimal pruned tree is typically chosen. The estimation of the one-step deviance for the pruned trees is performed by V-fold cross-validation; thus, the complexity parameter α that minimizes the average cross-validated deviance residuals over the V cross-validation subsamples is chosen.

Finally, once the optimal pruned tree has been chosen, the full-likelihood estimation procedure is used for obtaining the estimates of relative risk θ_k at each terminal node $\mathrm{W_k}$, repeating the two following iterative steps until convergence. Firstly, the Nelson–Aalen estimate (Nelson 1969) of baseline cumulative hazard at the u^{th} iteration is obtained using the current estimates $\hat{\theta}_k^u$ of θ_k:

$$\hat{H}_0^u(t) = \sum_{i:\tau_i \le t} \frac{\delta_i}{\sum_{k \in \tilde{\mathscr{T}}} \sum_{i:\tau_i \ge t;\, i \in W_k} \hat{\theta}_k^u} \,. \tag{2}$$

Next, the estimate $\hat{\theta}_k^{u+1}$ of θ_k is evaluated, using the current estimate $\hat{H}_0^u(t_i)$:

$$\hat{\theta}_k^{u+1} = \frac{\sum_{i \in W_k} \delta_i}{\sum_{i \in W_k} \hat{H}_0^u(t_i)} \,. \tag{3}$$

The final output is a tree with a relative risk (measure of risk with respect to the root node) estimated at each terminal node.

Further details can be found in LeBlanc and Crowley (1992).

2.1.2 Conditional inference trees

Conditional Inference Trees (Hothorn et al. 2006) have been introduced to address two main issues of other recursive partitioning algorithms: (*i*) overfitting and (*ii*) selection bias toward covariates with many possible splits or missing values. To this extent, CITs stand in the conditional inference framework and are based on the theory of permutation tests developed by Strasser and Weber (1999).

The basic idea of this algorithm is dividing the variable and split point selection in two distinct steps.

At each node, the choice of the splitting variable is made after testing the global null hypothesis of independence between any of the L covariates and the response variable. This global null hypothesis is seen as the combination of L partial null hypotheses testing the independence between the outcome and a specific covariate X_l ($l = 1, ..., L$). Once each partial null hypothesis is tested, a global test is carried out, and the global null hypothesis is rejected if the minimum adjusted p-value (e.g., Bonferroni-adjusted p-values or a minimum p-value resampling approach) is less than the significance level α. If this hypothesis cannot be rejected, the covariate X_l with the strongest association to the outcome, i.e., the one with the minimum p-value, is selected.

In survival analysis, the linear statistic adopted for the hypothesis testing can use the log-rank score or Savage scores as influence function (Hothorn et al. 2006).

Once a given splitting variable is chosen, the split can be defined using the same permutation test framework. The goodness of each split is evaluated on the basis of a two sample linear statistic that measures the discrepancy between the two candidate children nodes, and the split maximizing this test statistic is chosen.

The algorithm stops when the global null hypothesis of independence between the outcome and any of the covariates cannot be rejected at the given significance level α.

Afterward, once the final tree is obtained, for each terminal node the Kaplan–Meier curve (Collett 2015) is estimated and plotted.

The parameter α plays an important role in the model and can be interpreted in two ways: as a prespecified nominal level of the association tests or as a hyperparameter linked to the tree size.

Further details about the test statistics used can be found in Hothorn et al. (2006).

2.2 Random survival forests

Random Survival Forests (RSF) are an extension of random forests (Breiman 2001) to censored data. One of the most used algorithms is the one proposed by Ishwaran et al. (2008). This method follows the prescriptions laid out by Breiman (2003) and it basically works as classical random forests. It starts with B bootstrapped samples drawn with replacement from the original one; then, at each node a random subset of covariates is considered for splitting and B full-grown trees are obtained. In particular, trees can be built using different splitting criteria aimed to maximize the survival difference between children nodes and under the constraint that each node contains at least a given number of events. Possible choices include:

- a log-rank splitting rule based on the maximization of the log-rank test statistic (Kleinbaum and Klein 2012);
- a conservation-of-events splitting rule that splits nodes by finding children nodes closest to the conservation-of-events principle (defined later on);
- a log-rank score rule that splits nodes using a standardized log-rank statistic;

- a random log-rank splitting rule that uses, among the p candidate variables, the one with maximum log-rank statistic evaluated at a random selected split point.

For each terminal node W_k the cumulative hazard function (CHF) is evaluated through the Nelson–Aalen estimator:

$$\hat{H}_k(t) = \sum_{t_k \le t} \frac{d_k}{n_k} \quad k \in \tilde{\mathcal{T}},$$

where d_k and n_k are, respectively, the number of events and of subjects at risk at time t_k in the node W_k. Thus, an observation with covariates $\boldsymbol{x}_i$ falling in the terminal node W_k has a CHF $H(t|\boldsymbol{x}_i) = \hat{H}_k(t)$.

Finally, once the B trees are grown, an ensemble estimate can be obtained averaging all the obtained estimates. In particular, a bootstrap estimate and an out-of-bag (OOB) estimate can be defined. The first one is obtained averaging the CHFs of all the B trees; the second one is, instead, an average over the trees in which the corresponding observation was not used for growing them.

Under fairly general conditions, the conservation-of-events principle holds, i.e., the sum of the CHF over the observed times (both censored and not) equals the total number of events (Ishwaran et al. 2008).

An alternative predicted outcome to the CHF is the so-called "*mortality*", a measure that can be interpreted in terms of expected number of events under a null hypothesis of similar survival behavior (Ishwaran et al. 2008):

$$M_i = \mathbb{E}_i\left(\sum_{s=1}^{N} H(T_s|\boldsymbol{x}_i)\right)$$

where $\mathbb{E}_i$ is the expectation under the null hypothesis that all s are similar to i. Even in this case, both the bootstrap and the OOB ensemble estimates can be obtained.

One of the main advantages of using random survival forests instead of survival trees, is their stability with respect to noise data (Hastie et al. 2009). Moreover, as survival trees, they also allow to analyze both linear and nonlinear relationships, and to detect interactions between covariates without the need of specifying them beforehand, differently from Cox (Cox 1972) and Cox penalized models (Tibshirani 1997).

2.3 Variable importance

Variable importance (VIMP) is a measure used to evaluate the predictive power of each covariate X_l in predicting the outcome of interest. Different methods are available, according to the model used.

Relative Risk trees provide an overall measure of VIMP by summing the goodness of split measures for each split for which it was the primary variable, plus the adjusted agreement for all splits in which it was used as a surrogate (Therneau et al. 2022).

For Conditional Inference Trees, instead, VIMP can be assessed through the permutation approach, proposed by Breiman (2001), or through its conditional version (Strobl et al. 2008). In particular, the first method is based on the introduction of noise into the data by randomly permuting the values of X_l after the tree has been grown. The permuted data are then run down the corresponding tree, and the predictions for each observation are recorded. Variable importance is then evaluated as the increase or decrease in error compared to the error measured when no variables were permuted. Higher values indicate more important variables.

However, it has been shown that this approach overestimates the importance of correlated predictor variables (Strobl et al. 2008). To address this issue, the conditional approach proposed by Strobl et al. (2008) permutes X_l only within groups of observations with the same value of the covariates, to preserve their original correlation structure.

Finally, for Random Survival Forests there exist many different approaches for evaluating VIMP. Besides the permutation approach, the ”anti” and the ”random” methods (Ishwaran et al. 2021) can be used. With the first one, the OOB observations are dropped down their in-bag survival tree and, whenever a split for X_l is encountered, the observation is assigned to the opposite split. The ”random” approach instead, assigns the OOB observations randomly to the left/right node when X_l is used for the split.

In the end, for both methods, the CHF estimate from each of these trees is evaluated and the ensemble estimate is obtained. Finally, the VIMP for X_l is obtained evaluating the difference between the prediction errors for the original ensemble and for the new ensemble estimate obtained with the wrong split assignment. Large importance values denote variables with predictive ability, whereas zero or negative values identify non-predictive variables.

An alternative approach for evaluating VIMP for RSF is related to the notions of *maximal l-subtree*, i.e., the largest subtree whose root node is split using X_l, and *minimal depth*, that is the shortest distance from the root node to the parent node of the maximal l-subtree. A smaller minimal depth indicates that X_l has a greater impact on prediction. The strongest variables are then determined using a threshold based on the mean of the minimal depth distribution (Ishwaran et al. 2010).

Therefore, variable importance is a useful tool for assessing the predictive contribution of each covariate; however, it is important to also consider some potential drawbacks. In particular, when predictors are correlated, VIMP may be overestimated (Ishwaran et al. 2008). Furthermore, VIMP can be influenced by factors such as the number of categories and the scale of measurement of the predictor variables, which do not necessarily reflect the true importance of those variables (Strobl et al. 2007).

2.4 Performance evaluation measures

Different evaluation metrics exist for assessing the performance of survival models. In this study, models’ performance has been assessed through the prediction error rate and the time-dependent Area Under the Curves (AUCs), two discrimination metrics that allow to evaluate the model's ability to distinguish between low and high risk observations.

The prediction error rate is evaluated as the complement to one of the Harrell Concordance Index (C-index) C_H (Harrell et al. 1982). This index is a rank-order statistic that measures the proportion of all comparable pairs of observations for which the predicted and actual outcomes are concordant. The pair is concordant if the first observation has a higher actual survival time and a better survival prediction than the second one. Then, the prediction error rate is equal to

$$PE = 1 - C_H = 1 - \frac{1}{C} \sum_{t:\delta_i=1} \sum_{j:t_i<t_j} I[\hat{S}(t_i) < \hat{S}(t_j)] \ .$$

where $I[\cdot]$ is the indicator function, δ_i is the binary event indicator, t_i is the observed time for the i^{th} observation, $\hat{S}(T_i)$ is the estimated survival function for the i^{th} observation and C is the overall number of comparable pairs, i.e., the number of pairs where the censoring occurs later than the event time.

The index ranges from 0 to 1, with a value of 0.5 indicating that the prediction is equal to a random guessing.

Time-dependent AUCs are, instead, an extension of classical AUCs obtained evaluating the area under the ROC curve at different timepoints. The main differences from classical ROC analysis are the presence of censoring and the fact that the outcome of an observation can change over time (Park et al. 2021).

Different approaches can be used on the basis of the definition of events and non-events (Park et al. 2021). In this case study the *cumulative/dynamic* approach has been used; at a given timepoint t, events are defined as all observations who experienced the event between the interval $[0, t)$, while non-events are all those observations who were event-free at time t. According to these definitions, sensitivity, specificity and the AUC can be evaluated (Kamarudin et al. 2017).

Finally, the time-dependent AUC allows to evaluate the model's ability of discriminating the binary outcome (event/no event) at each time point. Values closer to 1 indicate a better performance.

In the case study described in the following sections, analyses have been carried out using R (R Core Team 2021). In particular, the `rpart` (Therneau et al. 2022), `partykit` (Hothorn and Zeileis 2015), and `randomForestSRC` (Ishwaran et al. 2022) packages have been used for fitting RRTs, CITs, and RSFs, respectively. Finally, `riskRegression` (Gerds et al. 2015) and `Hmisc` (Harrell Jr and Dupont 2006) have been used for evaluating the time-dependent AUCs and the C-index.

3 Data and study design

The 2020-2021 NBA regular season was examined. In detail, the season was split in two segments with respect to the All-Star Game: pre- and post-AS. The first part of the season was used for extracting the baseline covariates just before the AS. The second part was instead used as follow-up period. Baseline covariates were obtained from the NBA website; in addition, information about the selection for the AS and the participation to the NBA G-League were included. Play-by-play data, kindly made available by BigDataBall (www.bigdataball.com, a reliable source of validated and verified data for NBA), have been used for extracting the survival outcome. In order to make the number of achievements comparable among players, all the variables indicating the total number of achievements of each player have been normalized dividing them by his minutes played. The list of all the variables used in this study can be found in Table 1, together with the corresponding mean and standard deviation (SD).

In detail, let $j = 1, 2, \dots, J$ be the ordered sequence of made shots of a given player, and m_{ij} be the amount of minutes played by the player i when his j^{th} made shot was recorded. Note that this involves treating the time variable as playing-time, instead of clock-time, so that it increases when the player is in the court and remains constant when he is not playing. Then, the cumulative amount of points gained by the player i until his j^{th} made shot is denoted by p_{ij}. The event of interest for

Table 1 List of variables collected before the follow-up period (during the pre-All-Star season segment) and corresponding descriptive statistics

Variables	Mean (SD)
3PA - Three-point shots attempted	0.140 (0.078)
3P% - Percentage of three-point shots made *	32.062 (13.561)
2PA - Two-point shots attempted	0.207 (0.101)
2P% - Percentage of two-point shots made *	50.718 (12.524)
FTA - Free throws attempted	0.085 (0.060)
FT% - Percentage of free throws made *	72.760 (20.128)
OREB - Offensive rebounds	0.045 (0.041)
DREB - Defensive rebounds	0.142 (0.060)
AST - Assists	0.093 (0.061)
TOV - Turnovers	0.054 (0.029)
STL - Steals	0.031 (0.015)
BLK - Blocks	0.022 (0.024)
PF - Personal fouls	0.089 (0.042)
DD2 - Double doubles	0.003 (0.005)
TD3 - Triple doubles	0.0002 (0.0009)
$+/-$ - Plus/minus	−0.010 (0.248)
Age *	25.816 (4.098)
Percentage won matches (player) *	13.131 (6.305)
MIN - minutes played *	605.618 (365.542)
Percentage won matches (team - per game) *	0.506 (0.126)
PTS-T - points gained by the team (per game) *	112.295 (3.893)
All-Star Game	N.A
NBA G-League	N.A

Mean and Standard Deviations (SD) have been provided for all the normalized and non-normalized (denoted with *) quantitative variables. Normalized variables have been obtained by dividing the observed values by the minutes played by each player. N.A.: Not Applicable evaluating the survival probability is defined as the exceeding of a given threshold P of scored points. Different thresholds have been examined, according to different percentages of censoring, ranging from 10% to 75%.

Let

$$j^*(P) = \text{argmin}_{j=1,\dots,J}\ p_{ij} > P$$

be the made shot, if it exists, that leads the i^{th} player to exceed for the first time the threshold P. So, $m_{ij^*(P)}$ is the amount of minutes played by the player when exceeding the threshold.

The outcome of the study is then composed of

- the event indicator $\delta_i = I[j^*(P)]$, taking value 1 if $j^*(P)$ exists and 0 otherwise (i.e., if the player never exceeds the threshold P and, consequently, ends the study as a censored subject);
- the time-to-event

$$\tau_i = \begin{cases} m_{ij^*(P)} & \text{if } \delta_i = 1 \\ m_i & \text{if } \delta_i = 0 \end{cases} \tag{4}$$

where m_i denotes the total minutes played by the player i in the analyzed season segment.

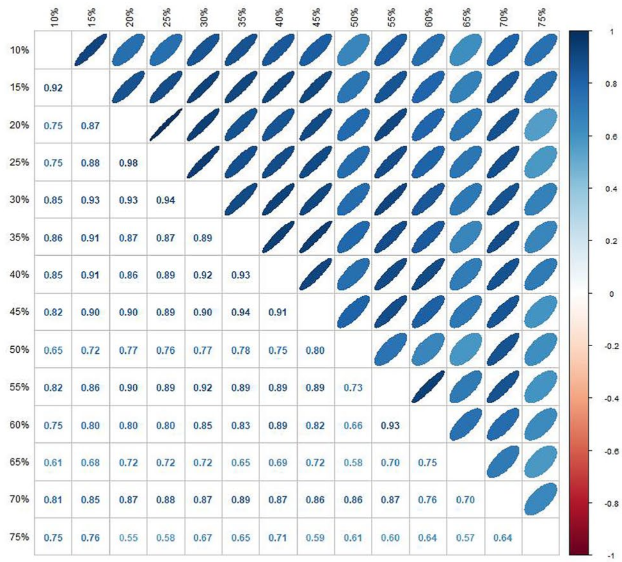
4 Results

The overall sample consisted of 359 NBA players, after having excluded those who played less than 48 minutes during the post-AS season segment and those who changed team during the season.

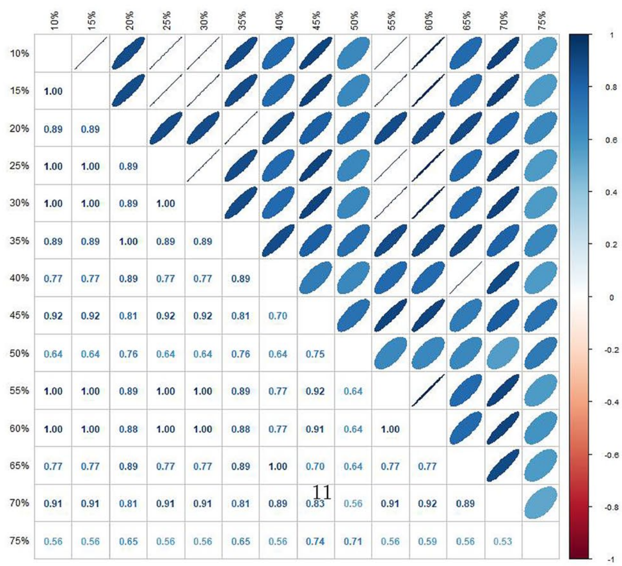
Firstly, different thresholds have been analyzed according to different percentages of censoring. In particular, percentages ranging from 10% to 75% (with a step of 5%) of censoring in the sample have been considered. Survival trees (RRTs and CITs) and RSFs have been fitted in the different settings, and measures of variable importance have been extracted for each algorithm. Then, for each method, the agreement between the variable importance measures obtained in the different frameworks has been evaluated using the Spearman coefficient ρ. Figure 1 shows that the results obtained with all the three machine learning algorithms have a high agreement through the different settings. For this reason, and for ensuring comparability with the results obtained in Macis et al. (2023a, 2023b), only the results related to a 50% of censoring in the sample will be shown hereafter.

It is also worth noting that, overall, RSFs exhibit a higher agreement between the variables importance measures through the different censoring frameworks. This is partly a consequence of the higher stability of RSFs with respect to trees.

The threshold P ensuring a 50% of censoring in the sample was equal to 255 points (median of the points scored by the players in the sample). First of all, an RRT was fitted and pruned through the cost-complexity algorithm (Fig. 2). Each terminal

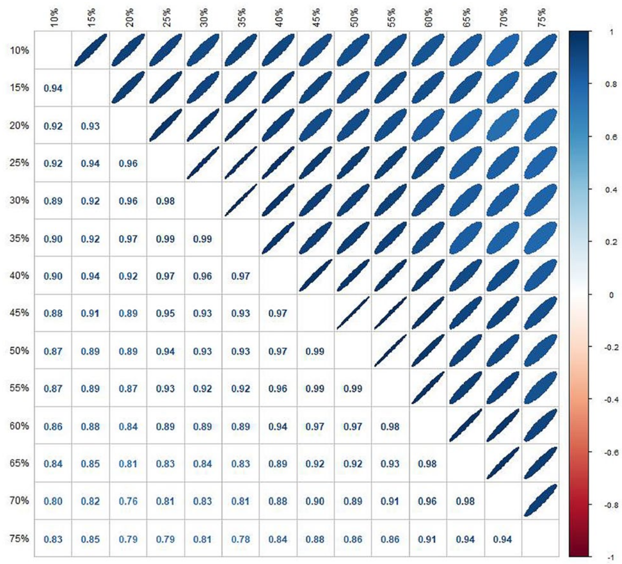


(a) Agreement of the variables importance measures obtained fitting Relative Risk Trees



(b) Agreement of the variables importance measures obtained fitting Conditional Inference Trees

Fig. 1 Agreement of the variables importance measures obtained fitting Relative Risk Trees (**a**), Conditional Inference Trees (**b**), and Random Survival Forests (**c**) in different settings, according to the observed amount of censoring in the sample



(c) Agreement of the variables importance measures obtained fitting Random Survival Forests

Fig. 1 (continued)

node reports the following information: (i) the estimated measure of relative risk; (ii) the ratio between the number of events and observations; (iii) the percentage of subjects in the node. The first splitting variable is the number of two-point shots attempted (2PA), followed by the number of three-point shots attempted (3PA), and the percentage of realized three-point shots (3P%). The other selected variables are the number of free throws attempted (FTA), the percentage of realized free throws (FT%), and the selection for the AS. It can be noticed that the RRT allows to identify some interactions between variables. For example, it resulted that players who attempted a higher number of two-point shots (more than 28 per 100 minutes), realized more than 33% of their three-point shots in the first part of the season, and were additionally selected for playing in the AS, have the highest probability of exceeding the fixed points threshold earlier during the follow-up compared to other players (as pointed out by the measure of relative risk of the right terminal node, equal to 8.1). Moreover, as the number of attempted and realized shots decreases, the probability of scoring more than 255 points in a shorter time decreases substantially.

Similarly, Fig. 3 shows the final output obtained through the *ctree* algorithm (CIT). In each terminal node the corresponding survival curve is displayed, showing the probability of scoring more than 255 points later than a certain timepoint. The obtained tree has a different structure than that of the RRT; however, some of the identified variables are the same (2PA, 3PA, 3P%, and AS). In particular, also in this case, the first split is due to the number of 2PA. Differently, the free throws do not appear in the model. Finally, the amount of minutes played (MIN) and

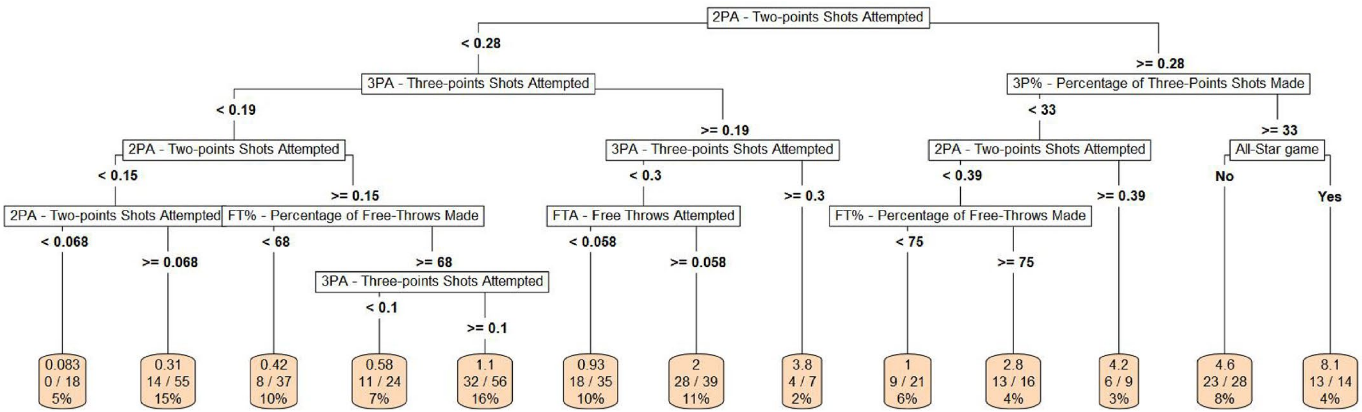


Fig. 2 Estimated Relative Risk Tree. Each leaf shows (1) the corresponding measure of relative risk, (2) the ratio between the number of events and observations in the node, and (3) the percentage of subjects in the node

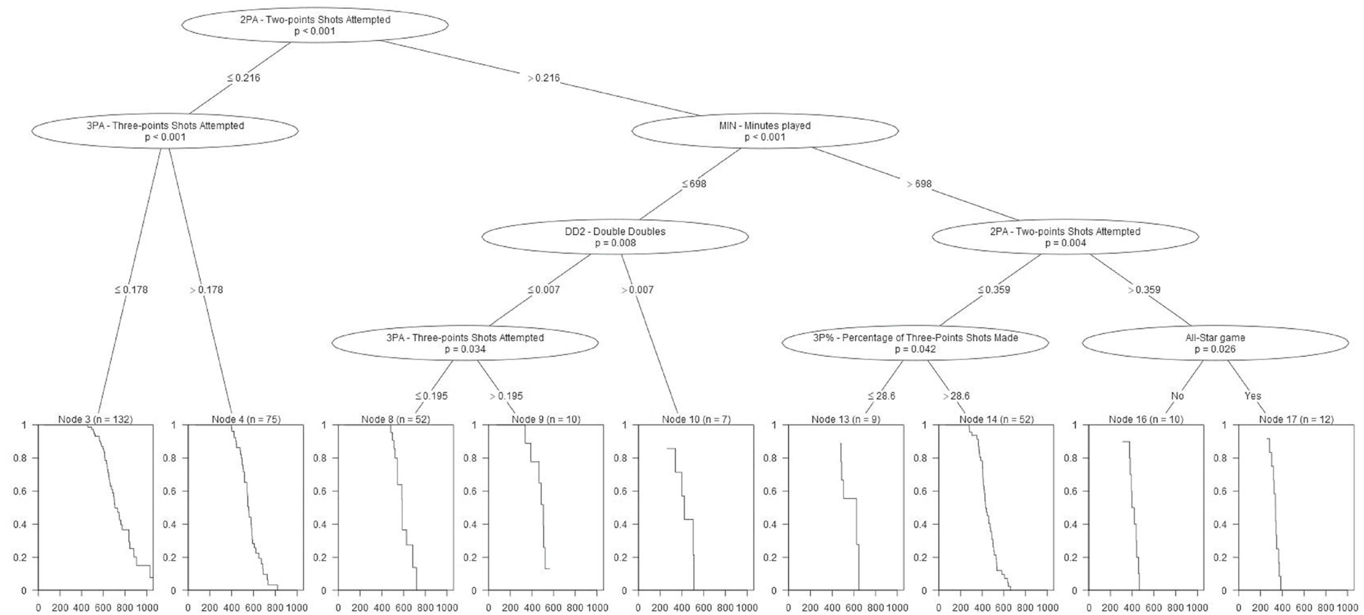


Fig. 3 Estimated Conditional Inference Tree. Each leaf represents the survival probability of the corresponding node, measuring the probability of not exceeding the threshold

double doubles (DD2) emerge as variables significantly related to the outcome. The obtained tree highlights, for example, that players selected for the AS who attempted more two-point shots and played more minutes in the first part of the season have a survival probability that decreases quickly (as shown by the Kaplan–Meier survival curve plotted in the corresponding leaf) and, therefore, have a higher probability of experiencing the event of interest in a shorter time.

However, it is known that the structure of trees is not stable with respect to little variations in the data used to grow the tree (Hastie et al. 2009). Instability of the trees is also confirmed by Fig. 9, which shows the RRT and CIT obtained on the basis of a slightly perturbed dataset (using 90% of the entire sample, selected at random); the resulting trees have a different structure, and some variables that had been previously selected, as the selection for AS, disappear.

Therefore, it has been preferred to resort to random survival forests, that are well known to be more stable than trees. In detail, an extensive grid search has been performed, focusing on the *mtry* and *nodesize* parameters, with *mtry* values ranging from 2 to 10, and *nodesize* values ranging from 5 to 20. Furthermore, various splitting methods (log-rank, bs.gradient, and log-rank score) were also tested. The RSF was then grown using 5000 trees, randomly selecting 10 variables at each split, with a *nodesize* of 5, and the log-rank splitting rule. In addition, to have a more interpretable representation of the RSF, the surrogate tree of the forest (Molnar 2025) was grown by using the obtained OOB predictions of "*mortality*", that measure the expected number of events under a null hypothesis of similar survival behavior (Ishwaran et al. 2008). Specifically, the extracted predictions were used as outcome variable of a regression tree. The final resulting tree (Fig. 4) does not need pruning, already being the optimal pruned tree. It can be observed that the only variables appearing in the tree are the number of attempted shots (in order of importance 2PA, 3PA, and FTA), and the selection for AS. Each terminal node shows a mean estimate of the "*mortality*" in that node. Also the RSF makes it possible to identify some interactions between the involved variables. For example, players who have attempted many two-point shots in the first part of the season and have been selected for AS have the highest expected cumulative probability of exceeding the threshold. On the contrary, players who have attempted a lower number of two- and three- point shots in the first season segment have the lowest expected number of events.

Finally, the variable importance measures, that give an idea of the impact of each predictor on the outcome, also taking into account the existing interactions and nonlinearities, have been computed for each algorithm. The obtained results are reported in Fig. 5.

For the CIT, the conditional VIMP has been evaluated using as error measure the integrated Brier Score (IBS) (Graf et al. 1999), a calibration measure that allows to quantify the accuracy of predicted survival probabilities over time.

It is important to notice that CITs assign a non-zero variable importance measure only to the covariates included in the final model. This occurs because any excluded covariates were found not to be significantly associated with the outcome during the tree-building process. Therefore, these variables do not contribute meaningfully to the predictive power of the model and, if included in the

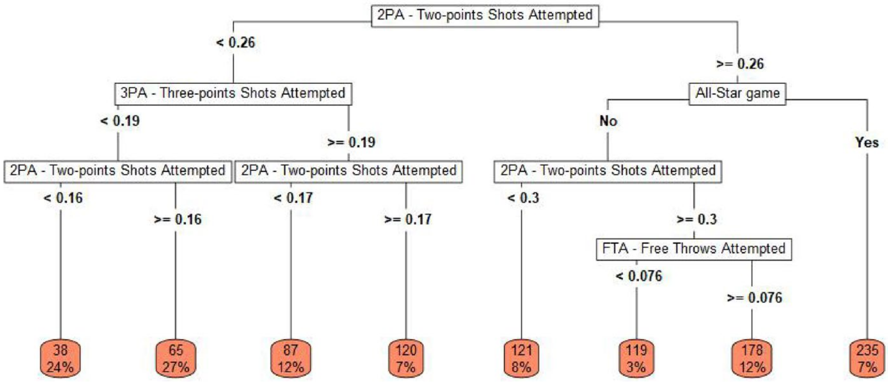


Fig. 4 Surrogate Survival Tree obtained using the out-of-bag predictions as outcome of a regression tree. In each terminal node the mean "*mortality*" is reported, together with the percentage of players included in that node

tree, would have likely resulted in non-informative splits (Sandri and Zuccolotto 2010).

For the RSF, instead, the "anti" method has been used, and VIMP has been evaluated on the basis of the prediction error rate.

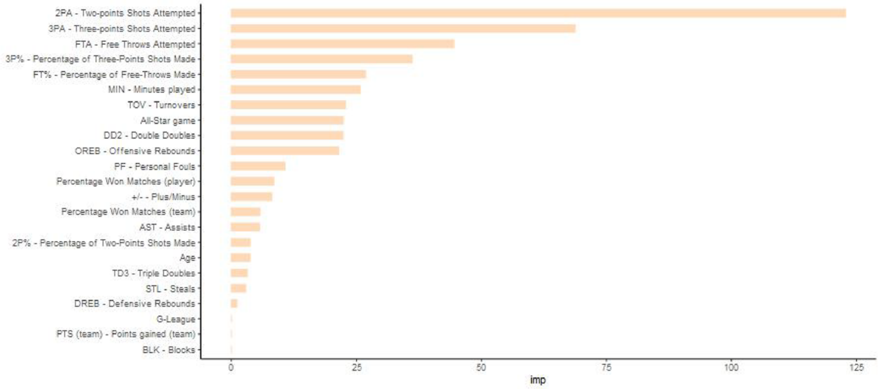
Even if each algorithm required different methods for assessing VIMP, leading to values that can be very different (as it can be seen on the x-axes of Fig. 5a, b and c), the results can still be compared in terms of variables' ranking.

The most important variables selected by the RSF (Fig. 5c) are the number of attempted two-point shots and the selection for AS, followed by the number of 3PA, DD2 and FTA. It is worth noting, that all these variables (with the only exception of DD2) are the ones that appear in the surrogate tree (Fig. 4). Similar results were also obtained using the maximal subtree approach.

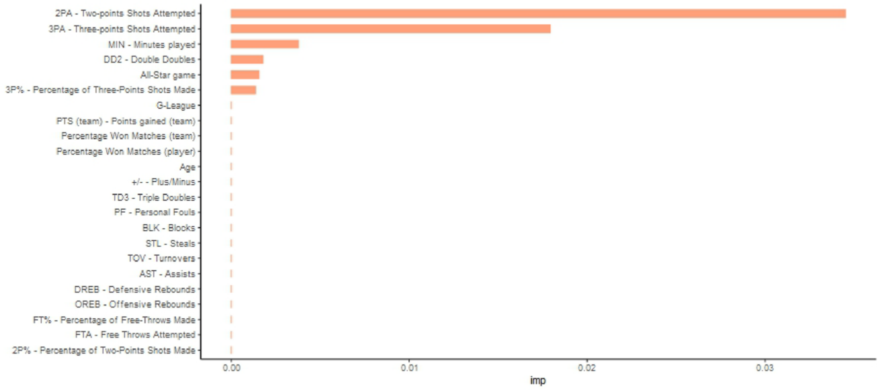
The variables' ranking obtained with the RSF is confirmed only in part by the rankings obtained from the RRT and the CIT (Fig. 5a and b, respectively). Indeed, it can be seen, for example, that the selection for AS has a lower importance in both the cases. Moreover, while the RSF suggests that the percentage of realized shots is not so important, the two survival trees give them higher importance.

The agreement of the variable importance measures obtained with the three algorithms has been evaluated, and the results highlighted a good agreement between both the two survival trees ($\rho = 0.64$), and between the RSF with the RRT and the CIT ($\rho = 0.69$ and $\rho = 0.66$, respectively).

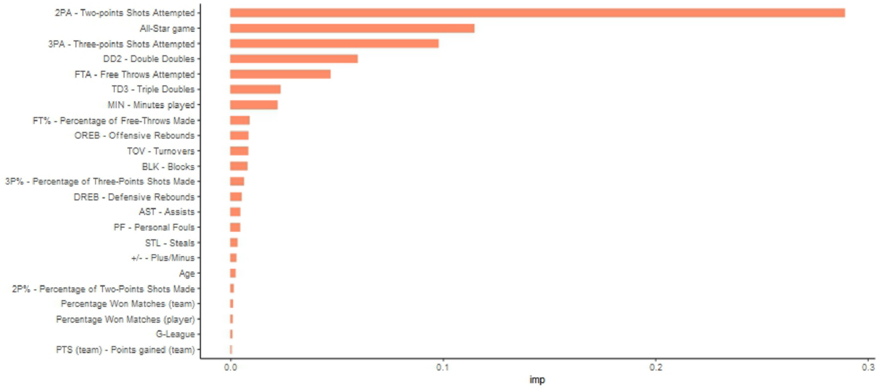
Finally, estimated survival probabilities at a given timepoint ($t = 480$), using the predictions from the RSF, have been plotted conditional on the three most important variables (2PA, 3PA and All-Star game), while all other quantitative variables have been set to their median values, and the NBA G-League has been set to "No". Figure 6 shows that as the number of 2PA and 3PA increases, the probability of scoring more than 255 points later than the examined timepoint (i.e., the survival probability) decreases. This effect is more pronounced with a higher number of two-point shots attempted (steeper curves). Moreover, it can be clearly observed that All-Star



(a) Relative Risk Tree



(b) Conditional Inference Tree



(c) Random Survival Forest

Fig. 5 Variable Importance measures obtained from Relative Risk Tree (**a**), Conditional Inference Tree (**b**), and Random Survival Forestx (**c**)

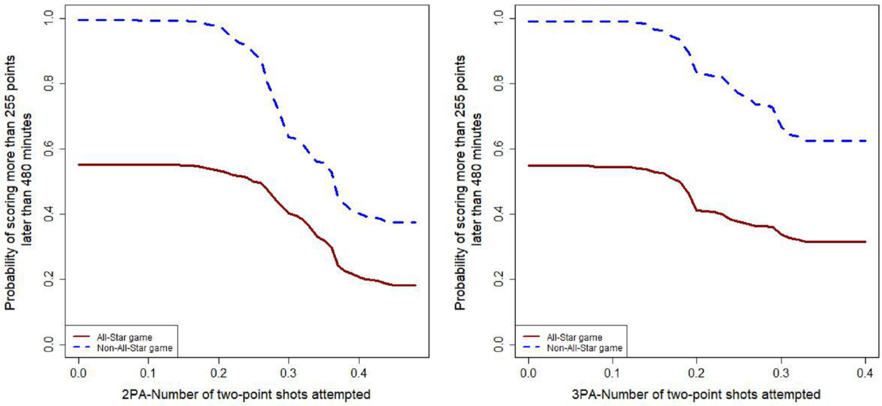


Fig. 6 Survival probabilities estimated using the predictions from the random survival forest at time $t = 480$ minutes, conditional on the All-Star game and on two- and three-point shots. All other variables have been set to their median values, and the NBA G-League has been set to “No"

game players have a lower probability of scoring more than 255 points later than 480 minutes regardless of the number of attempted shots, indicating a higher probability of exceeding the threshold earlier.

5 Performance evaluation

Figure 7 shows the out-of-bag prediction error rate of the RSF, i.e., the error evaluated using for each observation the corresponding OOB ensemble estimate. It can be seen that it becomes stable at a low value after around 1000 trees.

The performance of the models presented in this work, as well as those reported in Macis et al. (2023a), has been evaluated using the data from the two subsequent seasons, i.e., the 2021-2022 and 2022-2023 NBA seasons. It emerged that the best performing model was the Lasso Cox ($PE = 14.1\%$), followed by the Stepwise Cox ($PE = 14.6\%$) and the RSF ($PE = 16.2\%$). In contrast, the RRT and CIT had the highest prediction error rates ($PE = 21.5\%$ and $PE = 22.7\%$, respectively).

Finally, for comparing the performance ability of the three adopted machine learning algorithms (RRT, CIT, and RSF) with the two classical statistical models (Stepwise Cox and Lasso Cox), a sort of 10-fold cross-validation was performed using data from the three NBA seasons. During these three seasons 557 different players were observed, of whom 172 played in two seasons and 152 played in three seasons.

Due to the data structure and the presence of dependent observations (with players observed in more than one season), the use of a standard 10-fold cross-validation was not suitable. Therefore, a stratified sampling approach was applied. All players from the three seasons were considered, and 90% of these players were randomly assigned to the training set. Then, if one of these players appeared in more than one season, data from only one season was randomly selected for the training set,

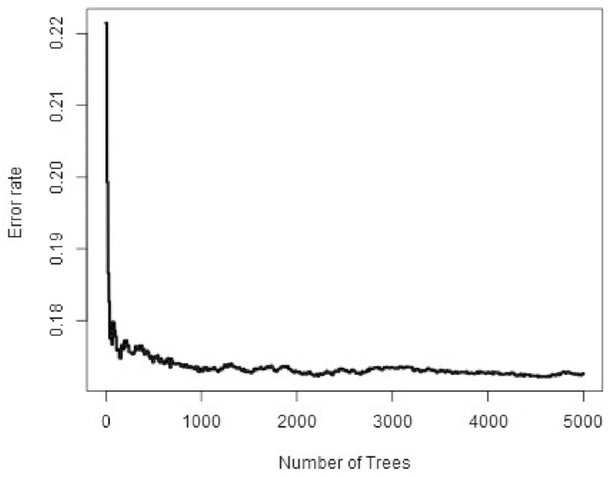


Fig. 7 Out-of-bag error rate of the random survival forest for number of trees

while the remaining observations were included in the test set. This procedure was repeated 10 times.

Figure 8 represents the estimated time-dependent AUCs. The results show that all the models have a good performance, with high values of AUC. In particular, it can be seen that the Lasso Cox and Cox model are the two models with the best predictive power, followed by the random survival forest, which has a similar performance to the two classical statistical models. Then, RRT and CIT are the two algorithms with the worst performance.

These results are also confirmed - in outline - by the overall prediction error rate (Table 2). It can be observed that the Lasso Cox model has the lowest prediction error rate, followed closely by Cox regression and RSF. Finally, RRTs and CITs have the highest error rates, even if, overall, the errors are quite low.

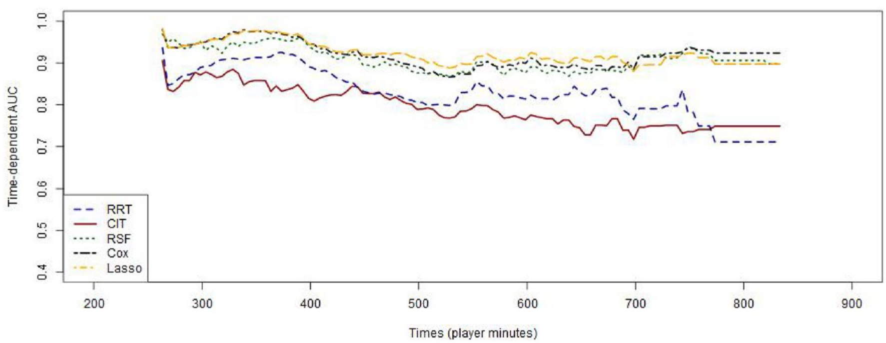


Fig. 8 Out-of-sample time-dependent AUCs evaluated for each fitted model. RRT: Relative Risk Tree; CIT: Conditional Inference Tree; RSF: Random Survival Forest; Cox: Stepwise Cox; Lasso: Lasso Cox

Table 2 Out-of-sample prediction error rate for each fitted model

Model	Error rate
Relative Risk Tree (RRT)	0.214
Conditional Inference Tree (CIT)	0.231
Random Survival Forest (RSF)	0.152
Cox regression	0.143
Lasso Cox	0.132

6 Discussion

Up to now, to the best of our knowledge, survival analysis has not been widely used for assessing basketball players' performance. This work aims to investigate the achievements most associated with the probability of exceeding a fixed amount of points P (in this work $P = 255$). Another study has already attempted to examine this problem using statistical models, in particular Cox regression and Lasso Cox (Macis et al. 2023a, b). In this study the authors showed that the most important variables were the number of two- and three-point shots, the selection for AS, and, to some extent, also the number of free throws, the number of double doubles, and the percentage of two-point shots. All these achievements increased the probability of exceeding the fixed threshold in a shorter time. Moreover, it seemed that also steals played a role in the exceeding of the threshold: a higher number of steals was associated with a lower probability of exceeding P. Interestingly, these results were stable across different percentages of censoring. However, in Macis et al. (2023a) interactions and possible nonlinear effects were not considered.

This study, therefore, aimed to take into account these aspects through the use of survival trees and random survival forests. Similarly to the results observed in Macis et al. (2023a), the results obtained by using these three machine learning algorithms highlighted a good stability through different settings defined by different percentages of censoring. For this reason, in this case study a single specific framework has been considered (50% of censoring). Furthermore, a good agreement was found between the variables importance measures obtained with the three algorithms.

The obtained results are coherent with the previous work, with the only exception concerning the role of steals that does not seem to have a high importance (see Fig. 5c).

In particular, one of the most interesting results is the utmost importance of the number of attempted two-point shots, rather than the three-point shots and the percentage of realized shots. Indeed, all the methods shown here and in Macis et al. (2023a) found that this variable was the most relevant in almost all the examined settings. Therefore, it seems that the most important thing is to attempt numerous baskets, and in particular two-point shots, rather than having a high success rate. Finally, the results shown here pointed that the defense variables do not have a high effect on the time needed by the player for exceeding the fixed threshold. However, differently from Macis et al. (2023a), machine learning algorithms also allowed to identify some interactions between 2PA and 3PA, as well as between 2PA, All-Star game selection, and FTA (Fig. 4). For example, players who attempted a higher number of two-point shots in the first part of the season and were selected for the All-Star game had a higher probability of scoring more than 255 points earlier in the second part of the season, compared to players who attempted a comparable number of two-point shots but were not selected for the AS.

From a methodological point of view, instead, the obtained results confirmed both the instability of trees and the power of random survival forests in overcoming this issue. Moreover, it emerged that the Lasso Cox and Cox regression had the best performance, and that also the RSF has a very good performance, similar to the other two best competitors. In the end, RRT and CIT resulted having the worst behavior, even if the overall prediction error rates were not very high. The similar performance of Lasso Cox, Stepwise Cox, and RSF suggests that sometimes simple traditional models may perform similarly or even better than complex methods as random survival forests, so one must not necessarily consider the approach with more complex methods as better. On the other hand, random survival forests allow an easier identification of interactions between variables and may offer a more structured interpretation of the results through the graphical representation of the surrogate survival tree. Therefore, the choice of the most appropriate model should be made by finding a balance between accuracy and interpretability, depending on the specific objective to be achieved.

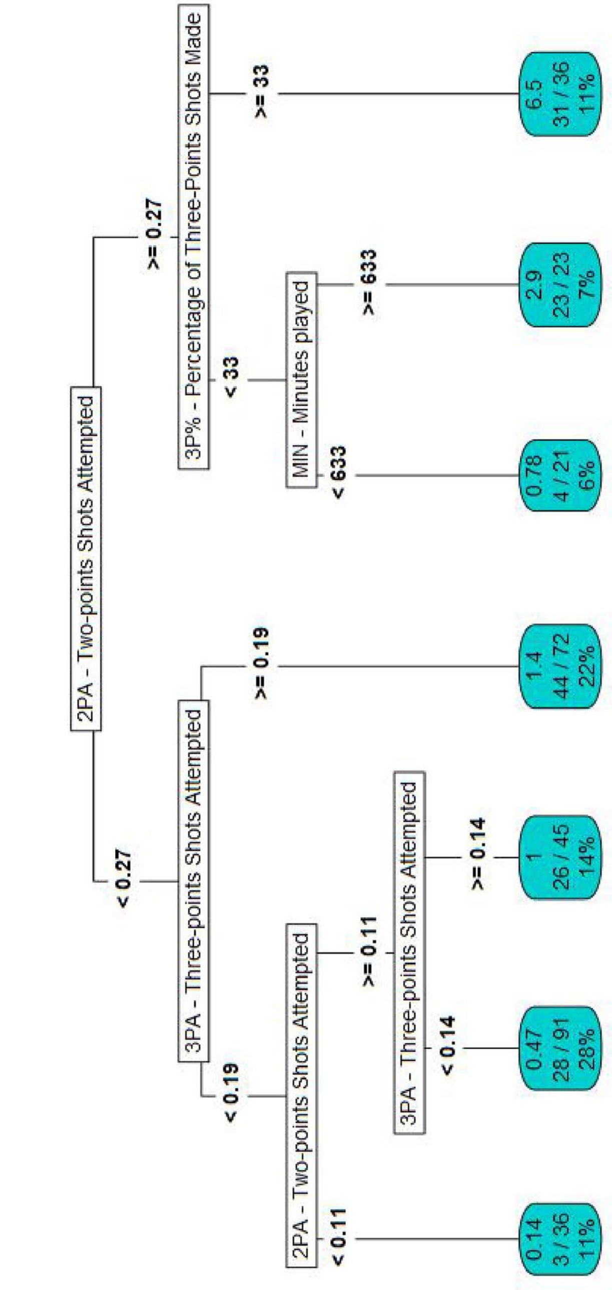
Concluding, the study has some limitations associated to possible issues related to the assumption of random censoring (Kleinbaum et al. 2012) that may be not respected (Macis et al. 2023a); however, the assumption of independent censoring (i.e., random censoring conditional on the set of covariates) can be retained valid (Kleinbaum et al. 2012; Macis et al. 2023a). Future research will try to definitely address this problem by using, for example, transformation models proposed by Hothorn and Zeileis (2021).

Moreover, a deeper analysis of variables interactions could be carried out, using, for example, maximal subtrees (Ishwaran et al. 2010).

Furthermore, considering that basketball players can be grouped according to different features (e.g., positions or teams), different models, such as frailty models, can take into account for this clustering effect through the inclusion of a random term called frailty. Examples of such models include regularized Cox frailty models (Groll et al. 2017), that can be implemented in the `PenCoxFrail` package (Groll 2016), and machine learning methods for clustered data (Hallett et al. 2014).

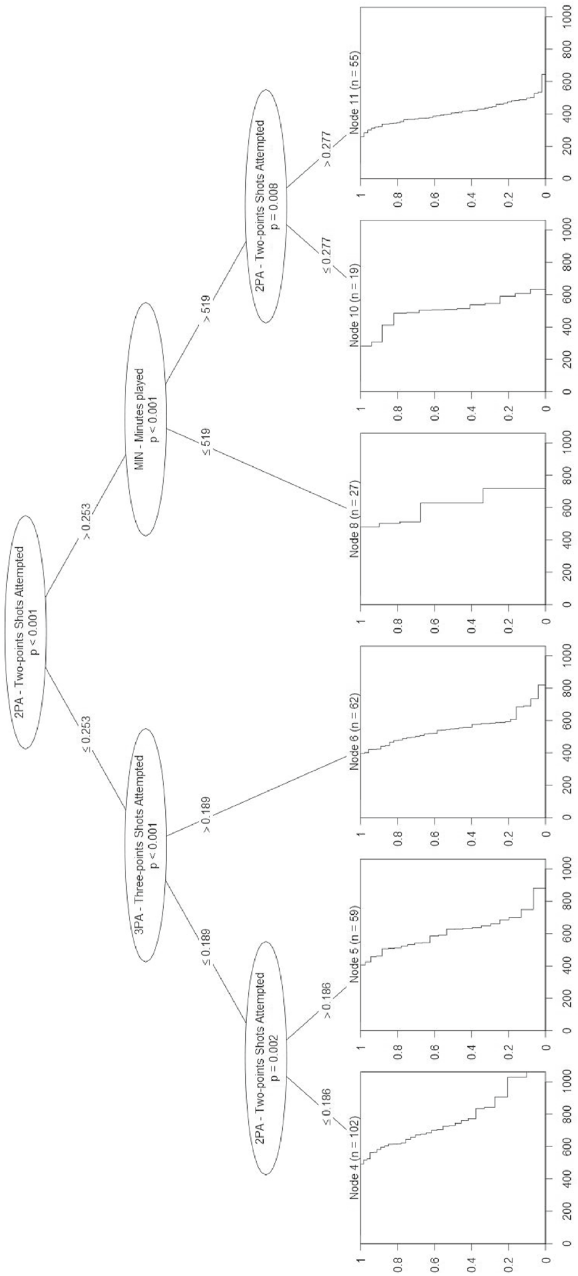
Finally, it could also be interesting to study possible time-varying effects of some covariates (e.g., number of shots attempted and percentage of shots realized) including time-varying coefficients into the models, as their effects may vary over time (Groll 2016).

Appendix A Trees instability



(a) Relative Risk Tree fitted on the perturbed dataset

Fig. 9 Survival Trees fitted on a perturbed dataset in which only 90% of the overall sample was included as training observations



(b) Conditional Inference Tree fitted on the perturbed dataset

Fig. 9 (continued)

Acknowledgements I would like to thank Prof. Paola Zuccolotto, Prof. Marica Manisera and Dr. Marco Sandri for their precious help in developing this research work. I would also like to thank the two anonymous reviewers for their valuable comments, which greatly improved the paper, and the Editor. The research was carried out in collaboration with the Big&Open Data Innovation Laboratory at University of Brescia (BODaI-Lab, https://bodai.unibs.it/), project "Big Data Analytics in Sport" (BDsports, https://bodai.unibs.it/bdsports/).

Declarations

Conflict of interest The author has no conflict of interest to disclose.

References

Back, F.A., Hino, A., Bojarski, W.G., et al.: Evening chronotype predicts dropout of physical exercise: a prospective analysis. Sport Sci. Health (2022). https://doi.org/10.1007/s11332-022-00963-8

Beynnon, B.D., Vacek, P.M., Murphy, D., et al.: First-time inversion ankle ligament trauma: the effects of sex, level of competition, and sport on the incidence of injury. Am. J. Sports Med. **33**(10), 1485–1491 (2005). https://doi.org/10.1177/036354650527549

Breiman, L.: Manual–Setting Up, Using, and Understanding Random Forests, v 4.0. Using_random_forests_V3 0 pdf URL ftp://ftp.stat.berkeley.edu/pub/users/breiman (2003)

Breiman, L.: Random forests. Mach. Learn. **45**(1), 5–32 (2001). https://doi.org/10.1023/A:1010933404324

Buist, I., Bredeweg, S.W., Bessem, B., et al.: Incidence and risk factors of running-related injuries during preparation for a 4-mile recreational running event. Br. J. Sports Med. **44**(8), 598–604 (2010). https://doi.org/10.1136/bjsm.2007.044677

Ciampi, A., Chang, C.H., Hogg, S., et al.: Recursive partition: A versatile method for exploratory-data analysis in biostatistics. In: Biostatistics. Springer, p 23–50 (1987)

Ciampi, A., Thiffault, J., Nakache, J.P., et al.: Stratification by stepwise regression, correspondence analysis and recursive partition: a comparison of three methods of analysis for survival data with covariates. Comput. Stat. Data Anal. **4**(3), 185–204 (1986)

Clarke, P.M., Walter, S.J., Hayen, A., et al.: Survival of the fittest: retrospective cohort study of the longevity of Olympic medallists in the modern era. BMJ (2012). https://doi.org/10.1136/bmj.e8308

Collett, D.: Modelling survival data in medical research. CRC Press (2015). https://doi.org/10.1201/b18041

Cox, D.R.: Regression Models and Life-Tables. Journal of the Royal Statistical Society Series B (Methodological) **34**(2), 187–220 (1972). https://doi.org/10.1111/j.2517-6161.1972.tb00899.x, URL http://www.jstor.org/stable/2985181

Csurilla, G., Fertő, I.: How long does a medal win last? Survival analysis of the duration of Olympic success. Appl. Econ. (2022). https://doi.org/10.1080/00036846.2022.2039370

Davis, R.B., Anderson, J.R.: Exponential survival trees. Stat. Med. **8**(8), 947–961 (1989)

Dekker, T.J., Godin, J.A., Dale, K.M., et al.: Return to sport after pediatric anterior cruciate ligament reconstruction and its effect on subsequent anterior cruciate ligament injury. JBJS **99**(11), 897–904 (2017). https://doi.org/10.2106/JBJS.16.00758

Del Corral, J., Barros, C.P., Prieto-Rodriguez, J.: The determinants of soccer player substitutions: A survival analysis of the Spanish soccer league. J. Sports Econ. **9**(2), 160–172 (2008). https://doi.org/10.1177/1527002507308309

Ekeland, A., Engebretsen, L., Fenstad, A.M., et al.: Similar risk of ACL graft revision for alpine skiers, football and handball players: the graft revision rate is influenced by age and graft choice. Br. J. Sports Med. **54**(1), 33–37 (2020). https://doi.org/10.1136/bjsports-2018-100020

Fynn, K.D., Sonnenschein, M.: An analysis of the career length of professional basketball players. Macalester Rev. **2**(2), 3 (2012)

Gerds, T.A., Scheike, T.H., Gerds, M.T.A.: Package 'riskregression', version 2020.12.08 (2015)

Gordon, L., Olshen, R.A.: Tree-structured survival analysis. Cancer Treat. Rep. **69**(10), 1065–1069 (1985)

Graf, E., Schmoor, C., Sauerbrei, W., et al.: Assessment and comparison of prognostic classification schemes for survival data. Stat. Med. **18**(17–18), 2529–2545 (1999). https://doi.org/10.1002/(SICI)1097-0258(19990915/30)18:17/18<2529::AID-SIM274>3.0.CO;2-5

Groll, A.: PenCoxFrail: Regularization in Cox Frailty Models. URL https://CRAN.R-project.org/package=PenCoxFrail, R package version 1.0.1 (2016)

Groll, A., Hastie, T., Tutz, G.: Selection of effects in cox frailty models by regularization methods. Biometrics **73**(3), 846–856 (2017). https://doi.org/10.1111/biom.12637

Gutiérrez, E., Lozano, S., González, J.R.: A recurrent-events survival analysis of the duration of Olympic records. IMA J. Manag. Math. **22**(2), 115–128 (2011). https://doi.org/10.1093/imaman/dpq005

Hallett, M., Fan, J., Su, X., et al.: Random forest and variable importance rankings for correlated survival data, with applications to tooth loss. Stat. Model. **14**(6), 523–547 (2014). https://doi.org/10.1177/1471082X14535517

Harrell, F.E., Jr., Dupont, M.C.: The Hmisc Package. R Package Vers. **3**(0–12), 3 (2006)

Harrell, F.E., Califf, R.M., Pryor, D.B., et al.: Evaluating the yield of medical tests. JAMA **247**(18), 2543–2546 (1982). https://doi.org/10.1001/jama.1982.03320430047030

Hastie, T., Tibshirani, R., Friedman, J.H., et al.: The elements of statistical learning: data mining, inference, and prediction, vol 2. Springer (2009). https://doi.org/10.1007/978-0-387-21606-5

Hopkins, W.G., Marshall, S.W., Quarrie, K.L., et al.: Risk factors and risk statistics for sports injuries. Clin. J. Sport Med. **17**(3), 208–210 (2007). https://doi.org/10.1097/JSM.0b013e3180592a68

Hothorn, T., Zeileis, A.: Partykit: A modular toolkit for recursive partytioning in R. J. Mach. Learn. Res. **16**(1), 3905–3909 (2015)

Hothorn, T., Zeileis, A.: Predictive distribution modeling using transformation forests. J. Comput. Graph. Stat. **30**(4), 1181–1196 (2021). https://doi.org/10.1080/10618600.2021.1872581

Hothorn, T., Hornik, K., Zeileis, A.: Unbiased recursive partitioning: a conditional inference framework. J. Comput. Graph. Stat. **15**(3), 651–674 (2006). https://doi.org/10.1198/106186006X133933

Howell, D.R., Potter, M.N., Kirkwood, M.W., et al.: Clinical predictors of symptom resolution for children and adolescents with sport-related concussion. J. Neurosurg. Pediatr. **24**(1), 54–61 (2019). https://doi.org/10.3171/2018.11.PEDS18626

Ishwaran, H., Kogalur, U.B., Kogalur, M.U.B.: Package 'randomforestsrc'. breast 6:1 (2022)

Ishwaran, H., Lu, M., Kogalur, U.B.: randomForestSRC: getting started with randomForestSRC vignette. http://randomforestsrc.org/articles/getstarted.html, URL http://randomforestsrc.org/articles/getstarted.html, [accessed date: 30/09/2024] (2021)

Ishwaran, H., Kogalur, U.B., Blackstone, E.H., et al.: Random survival forests. Ann. Appl. Stat. **2**(3), 841–860 (2008). https://doi.org/10.1214/08-AOAS169

Ishwaran, H., Kogalur, U.B., Gorodeski, E.Z., et al.: High-dimensional variable selection for survival data. J. Am. Stat. Assoc. **105**(489), 205–217 (2010). https://doi.org/10.1198/jasa.2009.tm08622

Jack, R.A., Sochacki, K.R., Hirase, T., et al.: Performance and return to sport after hip arthroscopic surgery in major league baseball players. Orthop. J. Sports Med. **7**(2), 2325967119825835 (2019). https://doi.org/10.1177/2325967119825835

Kamarudin, A.N., Cox, T., Kolamunnage-Dona, R.: Time-dependent ROC curve analysis in medical research: current methods and applications. BMC Med. Res. Methodol. **17**, 1–19 (2017). https://doi.org/10.1186/s12874-017-0332-6

Kleinbaum, D.G., Klein, M., et al.: Survival analysis: a self-learning text, vol 3. Springer (2012). https://doi.org/10.1007/978-1-4419-6646-9

Kleinbaum, D.G., Klein, M.: Kaplan-Meier Survival Curves and the Log-Rank Test", Springer New York, New York, NY, pp 55–96 (2012). https://doi.org/10.1007/978-1-4419-6646-9_2

Kontos, A.P., Elbin, R., Sufrinko, A., et al.: Recovery following sport-related concussion: integrating pre- and postinjury factors into multidisciplinary care. J. Head Trauma Rehabil. **34**(6), 394–401 (2019). https://doi.org/10.1097/HTR.0000000000000536

Kundu, M.G., Ghosh, S.: Survival trees based on heterogeneity in time-to-event and censoring distributions using parameter instability test. Stat. Anal. Data Min: The ASA Data Sci. J. **14**(5), 466–483 (2021)

Lawrence, D.W., Richards, D., Comper, P., et al.: Earlier time to aerobic exercise is associated with faster recovery following acute sport concussion. PLoS ONE **13**(4), e0196062 (2018). https://doi.org/10.1371/journal.pone.0196062

LeBlanc, M., Crowley, J.: Relative risk trees for censored survival data. Biometrics (1992). https://doi.org/10.2307/2532300

LeBlanc, M., Crowley, J.: Survival trees by goodness of split. J. Am. Stat. Assoc. **88**(422), 457–467 (1993)

Lu, Y., Jurgensmeier, K., Till, S.E., et al.: Early ACLR and risk and timing of secondary meniscal injury compared with delayed ACLR or nonoperative treatment: a time-to-event analysis using machine learning. Am. J. Sports Med. (2022). https://doi.org/10.1177/03635465221124258

Macis, A., Manisera, M., Sandri, M., et al.: A Survival Analysis Study to Discover Which Skills Determine a Higher Scoring in Basketball. Statistica Applicata - Italian J. Appl. Stat. (2023a) https://doi.org/10.26398/IJAS.0035-009

Macis, A., Manisera, M., Sandri, M., et al.: Which Achievements Are Associated With a Better Offensive Performance in NBA? A Survival Analysis Study. In: 13th World Congress of Performance Analysis of Sport (WCPAS2022) & the 13th International Symposium on Computer Science in Sport (IACSS2022), Springer, https://doi.org/10.1007/978-3-031-31772-9 (2023b)

Macis, A.: Statistical Models and Machine Learning for Survival Data Analysis. PhD thesis, Università degli Studi di Brescia (2023)

Macis, A.: The role of the frailty in the evaluation of injury risk factors for National Basketball Association players. Computational Statistics (2024). https://doi.org/10.1007/s00180-024-01556-4

Macis, A.: Survival trees: a pathway among features and open issues of the main R packages. Electron. J. Appl. Stat. Anal. (2022). https://doi.org/10.1285/i20705948v15n3p479

Mahmood, A., Ullah, S., Finch, C.: Application of survival models in sports injury prevention research: a systematic review. Br. J. Sports Med. **48**(7), 630–630 (2014). https://doi.org/10.1136/bjsports-2014-093494.190

Mai, H.T., Chun, D.S., Schneider, A.D., et al.: Performance-based outcomes after anterior cruciate ligament reconstruction in professional athletes differ between sports. Am. J. Sports Med. **45**(10), 2226–2232 (2017). https://doi.org/10.1177/0363546517704834

Molnar, C.: Interpretable Machine Learning, 3rd edn. URL https://christophm.github.io/interpretable-ml-book (2025)

Moulds, K., Abbott, S., Pion, J., et al.: Sink or swim? A survival analysis of sport dropout in Australian youth swimmers. Scandinavian J. Med. Sci. Sports **30**(11), 2222–2233 (2020). https://doi.org/10.1111/sms.13771

Nelson, W.: Hazard plotting for incomplete failure data. J. Qual. Technol. **1**(1), 27–52 (1969). https://doi.org/10.1080/00224065.1969.11980344

Nelson, L.D., Tarima, S., LaRoche, A.A., et al.: Preinjury somatization symptoms contribute to clinical recovery after sport-related concussion. Neurology **86**(20), 1856–1863 (2016). https://doi.org/10.1212/WNL.0000000000002679

Nevo, D., Ritov, Y.: Around the goal: examining the effect of the first goal on the second goal in soccer using survival analysis methods. J. Quantit. Anal. Sports **9**(2), 165–177 (2013). https://doi.org/10.1515/jqas-2012-0004

Park, S.Y., Park, J.E., Kim, H., et al.: Review of statistical methods for evaluating the performance of survival or other time-to-event prediction models (from conventional to deep learning approaches). Korean J. Radiol. **22**(10), 1697 (2021). https://doi.org/10.3348/kjr.2021.0223

Pion, J., Lenoir, M., Vandorpe, B., et al.: Talent in female gymnastics: a survival analysis based upon performance characteristics. Int. J. Sports Med. **94**(11), 935–940 (2015). https://doi.org/10.1055/s-0035-1548887

Pratas, J.M., Volossovitch, A., Carita, A.I.: The effect of performance indicators on the time the first goal is scored in football matches. Int. J. Perform. Anal. Sport **16**(1), 347–354 (2016). https://doi.org/10.1080/24748668.2016.11868891

R Core Team: R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, URL https://www.R-project.org/ (2021)
Sandri, M., Zuccolotto, P.: Analysis and correction of bias in total decrease in node impurity measures for tree-based algorithms. Stat. Comput. **20**, 393–407 (2010). https://doi.org/10.1007/s11222-009-9132-0
Segal, M.R.: Regression trees for censored data. Biometrics pp 35–47 (1988)
Smith, K.L., Weir, P.L.: An Examination of Relative Age and Athlete Dropout in Female Developmental Soccer. Sports **10**(5), 79 (2022). https://doi.org/10.3390/sports10050079
Sochacki, K.R., Jack, R.A., Hirase, T., et al.: Performance and return to sport after hip arthroscopy for femoracetabular impingement syndrome in National Hockey League players. J. Hip Preservat. Surg. **6**(3), 234–240 (2019). https://doi.org/10.1093/jhps/hnz030
Strasser, H., Weber, C.: On the asymptotic theory of permutation statistics. https://doi.org/10.57938/ff565ba0-aa64-4fe0-a158-86fd331bee78 (1999)
Strobl, C., Boulesteix, A.L., Zeileis, A., et al.: Bias in random forest variable importance measures: illustrations, sources and a solution. BMC Bioinfo. **8**, 1–21 (2007)
Strobl, C., Boulesteix, A.L., Kneib, T., et al.: Conditional variable importance for random forests. BMC Bioinfo. **9**, 1–11 (2008). https://doi.org/10.1186/1471-2105-9-307
Therneau, T.M., Atkinson, E.J., et al.: An introduction to recursive partitioning using the RPART routines. Tech. rep, Technical report Mayo Foundation (2022)
Thomas, A.C.: Inter-arrival times of goals in ice hockey. J. Quantit. Analy. Sports (2007). https://doi.org/10.2202/1559-0410.1064
Tibshirani, R.: The lasso method for variable selection in the Cox model. Stat. Med. **16**(4), 385–395 (1997). https://doi.org/10.1002/(SICI)1097-0258(19970228)16:4<385::AID-SIM380>3.0.CO;2-3
Tozetto, A.B., Carvalho, H.M., Rosa, R.S., et al.: Coach turnover in top professional Brazilian football championship: A multilevel survival analysis. Front. Psychol. **10**, 1246 (2019). https://doi.org/10.3389/fpsyg.2019.01246
Venturelli, M., Schena, F., Zanolla, L., et al.: Injury risk factors in young soccer players detected by a multivariate survival model. J. Sci. Med. Sport **14**(4), 293–298 (2011). https://doi.org/10.1016/j.jsams.2011.02.013
Wangrow, D.B., Schepker, D.J., Barker, V.L., III.: Power, performance, and expectations in the dismissal of NBA coaches: a survival analysis study. Sport Manag. Rev. **21**(4), 333–346 (2018). https://doi.org/10.1016/j.smr.2017.08.002
Zeileis, A., Hothorn, T., Hornik, K.: Model-based Recursive Partitioning. J. Comput. Graph. Stat. **17**(2), 492–514 (2008). https://doi.org/10.1198/106186008X319331
Zuccolotto, P., Manisera, M.: Basketball data science: with applications in R. CRC Press (2020). https://doi.org/10.1201/9780429470615
Zumeta-Olaskoaga, L., Weigert, M., Larruskain, J., et al.: Prediction of sports injuries in football: a recurrent time-to-event approach using regularized Cox models. AStA Adv. Stat. Anal. (2021). https://doi.org/10.1007/s10182-021-00428-2