

Article

Ranking Features on Psychological Dynamics of Cooperative Team Work through Bayesian Networks

Pilar Fuster-Parra ^{1,*}, Alex García-Mas ^{2,†}, Jaume Cantallops ^{3,†}, F. Javier Ponseti ^{3,†} and Yuhua Luo ^{1,†}

¹ Department of Mathematics and Computer Science, Universitat Illes Balears, Carr. Valldemossa km. 7.5, Palma de Mallorca, Spain; y.luo@uib.es

² Faculty of Psychology, Universitat Illes Balears, Carr. Valldemossa km. 7.5, Palma de Mallorca, Spain; alex.garcia@uib.es

³ Department of Pedagogy, Universitat Illes Balears, Carr. Valldemossa km. 7.5, Palma de Mallorca, Spain; jaume.cantallops@uib.es (J.C.); xponseti@uib.es (F.J.P.)

* Correspondence: pilar.fuster@uib.es; Tel.: +34-971-171386; Fax: +34-971-173003

† All authors contributed equally to this work.

Academic Editor: Karl Grammer

Received: 18 November 2015; Accepted: 3 May 2016; Published: 9 May 2016

Abstract: The aim of this study is to rank some features that characterize the psychological dynamics of cooperative team work in order to determine priorities for interventions and formation: *leading positive feedback*, *cooperative manager* and *collaborative manager* features. From a dataset of 20 cooperative sport teams (403 soccer players), the characteristics of the prototypical sports teams are studied using an average Bayesian network (BN) and two special types of BNs, the Bayesian classifiers: *naive Bayes* (NB) and *tree augmented naive Bayes* (TAN). BNs are selected as they are able to produce probability estimates rather than predictions. BN results show that the antecessors (the “top” features ranked) are the team members’ expectations and their attraction to the social aspects of the task. The main node is formed by the cooperative behaviors, the consequences ranked at the BN bottom (ratified by the TAN trees and the instantiations made), the roles assigned to the members and their *survival* inside the same team. These results should help managers to determine contents and priorities when they have to face team-building actions.

Keywords: Bayesian networks; cohesion team; collaborative work; cooperative teams; naive Bayes; symmetric independencies; team dynamics; tree augmented naive Bayes

1. Introduction

In high performance sports, it is crucial to rank features to determine their degree of influence in a cooperative work team, with the ultimate objective of maximizing the team’s performance. Along these lines, several studies have been carried out in the field of social psychology, where features, such as motivation, cohesion, leading positive feedback or social attraction, play an important role due to their close relation to performance [1–5].

Cooperative team work is a fact well known today [6], even when taking into account the associated psychological factors’ [7] analysis. However, it is difficult to find studies that relate the equipment they need to cooperate with effective performance [8]. Some psychological factors of collaborative teams have been studied more than others in terms of performance, such as cohesion [9,10] or group facilitators and blockers’ roles [11]. Features, such as the specificity of the jobs [12] (or playing positions, in this case), generating performance expectations, the motivational climate generated by the manager [13] and the leadership styles of the coach, which are not usually studied, are considered here in this paper.

Bayesian networks (BNs) [14–18], well suited to reason with uncertain domain knowledge, can be applied to aid teams by providing cooperative and collaborative work characterization estimates. BNs have been proven to be a strong tool to discover the relationships between variables that attempt to separate out indirect from direct association [8,19–21] and can capture the way an expert understands the relationships among all of the features [22]. They enable human experts to better understand the modeled domain. Contrary to deterministic understanding of the *causality* phenomenon [23], BN modeling lies within the data mining and machine learning literature [24,25]. The network structure is a directed acyclic graph (DAG) where each node represents a random variable [26,27], and the arcs may represent causality [28–30]. BNs combine graph theory and probability theory to represent relationships between variables (nodes in the graph) [8,19]. They give a compact representation of a joint probability distribution via conditional independence. The major advantage of BN is the ability to represent and, hence, understand knowledge. BNs are the best known classifiers that are able to provide the probability distributions concisely and comprehensibly [31,32]. In [33], the authors considered BN model with the naive Bayes algorithm as one of the most effective classification algorithms.

Recently, there has been increasing attention regarding the application of BNs in competitive sport contexts. They have been used to analyze team performance related to different types of cooperation between team members and various motivational climates generated by teams' coaches in sports performance [20,34], achieving different results from those obtained by classical statistics.

Based on the dataset obtained by our experimental study, the main objective of this work is to rank the features that characterize the *leading positive feedback* feature, the *cooperative manager* feature and the *collaborative manager* feature through an average BN, while comparing its performance with two special types of BNs, the Bayesian classifiers: *naive Bayes* (NB) and *tree augmented naive Bayes* (TAN). In the NB classifier, the class attribute is the single parent of each node of an NB network. The TAN classifier models relationships among the features of at most order one. The experimental results show that our average BN and the TAN models have similar performances on the classification of the *leading positive feature*, the *cooperative manager* feature and the *collaborative manager* feature, with the highest accuracy by the BN, 90%, 85% and 87%, respectively, closely followed by the TAN classifier, 90%, 84% and 87%, respectively.

The paper is organized as follows. Section 2 presents the materials and methods given the process to obtain the different cooperative team models. Section 3 presents the results of the *leading positive feedback*, *cooperative* and *collaborative manager* models through a BN. Section 4 presents some discussion. Finally, Section 5 concludes the paper.

2. Materials and Methods

2.1. Participants

The participants were male semi-professional football players from 20 teams who participated in the Third Division of a Spanish soccer League. A sample of 403 between the ages of 18 and 39 years with a median age of 24 years (IQR = 27–21.5) was used.

2.1.1. Procedure

The data were collected at the middle of the regular competitive season, when the majority of teams' dynamic processes are well engaged. All participants are legal adults, and their participation was voluntary; the authors obtained written consent from each participant. Participants completed the questionnaires in the changing room before a training session.

2.2. Instruments and Material

2.2.1. Leadership Behaviors

Coach leadership behaviors were assessed using an adapted Spanish version of the Leadership Sport Scale (LSS) [35]. This is a 40-item instrument designed to measure the following five dimensions of leadership: two of them measure decision-making style (democratic behaviors, autocratic behaviors); motivational tendencies are measured by the other two dimensions (social support and positive feedback). Additionally, the last one measures instruction behavior (training and instruction). Responses were rated on a 5-point scale ranging from strongly disagree (1) to strongly agree (5).

2.2.2. Group Environment Questionnaire

The Spanish version of the GEQ [36,37] was used to assess team cohesion. This inventory of 18 items comprises the following four factors: the group integration-task, which refers to an individual's beliefs about team closeness, similarity and bonding around the group's task; group integration-social, which refers to a team member's sense of group closeness, similarity and bonding as a social unit; individual attraction to the group-task, which refers to a group member's feelings about personal involvement in relation to shared group goals and productivity; and the individual attraction to the group-social, which refers to the team member's impressions of social interactions and personal acceptance within the group. These four factors are grouped into the two main factors of social cohesion and task cohesion. Responses were rated on a 5-point scale ranging from strongly disagree (1) to strongly agree (5).

2.2.3. Work Experience and Expectations

Work experience and expectations were measured through an *ad hoc* questionnaire containing the following questions, from "would have to play a lot less" to "would have to play a lot more".

1. *Work experience*: "How many times have you played in this team?"
2. *Expectations*: we asked players and managers to predict what position they believed they would occupy in the standings at the end of the season.

2.2.4. Specificity Workplace

The positions of the players on the pitch were divided into 4 groups: goalkeeper, defender, midfielder and striker.

2.3. Bayesian Networks Modeling

The BNs' representation is a DAG, of which the nodes are the random variables in our domain, and the edges may correspond to the direct influence of one node on another. They capture the dependencies and conditional independencies, which are represented by a joint probability distribution. BNs are not restricted to representing distributions satisfying the independence assumptions, which is implicit in the NB and TAN models.

To obtain a BN, it is necessary to determine a structure defined by a DAG and the conditional probabilities assigned to each node of the DAG. Therefore, to learn a BN implies two tasks: (i) *structural learning*, that is the identification of the topology of the BN; and (ii) *parametric learning*, that is the estimation of numerical parameters (conditional probabilities) given a network topology.

One important characteristic of BNs is the global Markov property, which states that any node X_i is conditionally independent of any other node given its *Markov blanket*, i.e., $I(X_i, non - Markov - blanket(X_i) | Markov - blanket(X_i))$; the *Markov blanket* of a node includes its parents, its children and the children's other parents (spouses). The global Markov property has been applied to select the features with influence on the *leading positive feedback*, *cooperative manager* and *collaborative manager* features.

An average BN and two specific classifiers among all of the different Bayesian classifiers, NB and TAN, have been selected to study the *leading positive feedback, cooperative manager* and *collaborative manager* features.

Learning a BN from data is a form of *unsupervised learning*, in the sense that the learner does not distinguish the class variable from the attribute variables in the data [38]. NB and TAN classifiers are special types of BN, where a *supervised learning* is performed.

A Bayes classifier assigns the most probable *a posteriori* (MAP) class to a given instance $\mathbf{x}_i = (x_{i1}, \dots, x_{in})$, as is shown in Equation (1):

$$\arg \max_j P(c_j | \mathbf{x}) = \arg \max_j P(\mathbf{x}, c_j), \quad (1)$$

which is optimal in terms of minimizing the conditional risk under a symmetric loss function [39], where $\mathbf{x}_i = (x_{i1}, \dots, x_{in})$ is an instance in a dataset.

BN classifiers [38] approximate $P(\mathbf{x}, c_j)$ with a factorization according to the DAG structure of a BN [40]. A Bayesian classifier applies Bayes' theorem (see Equation (2)):

$$P(c_j | \mathbf{x}_i) = \frac{P(\mathbf{x}_i | c_j) P(c_j)}{\sum_k P(\mathbf{x}_i | c_k) P(c_k)}. \quad (2)$$

The probability that a sample with characteristics \mathbf{x}_i belongs to a class c_j is denoted by $P(c_j | \mathbf{x}_i)$, *i.e.*, the posterior probability. The prior probability denoted by $P(c_j)$ is the probability that a sample belongs to a class c_j given no information on its characteristic values.

2.3.1. Conditional Independence of Triplets of Random Variables

Let us consider \mathbf{X} , \mathbf{Y} , \mathbf{Z} sets of random variables (features in the dataset). It is said that \mathbf{X} is conditionally independent of \mathbf{Y} given \mathbf{Z} in a distribution P , if P satisfies $P(\mathbf{X} | \mathbf{Y}, \mathbf{Z}) = P(\mathbf{X} | \mathbf{Z})$.

2.3.2. Bayesian Network Model

BN models estimate the joint probability distribution P over a vector of random variables $\mathbf{X} = (X_1, \dots, X_n)$. The joint probability distribution factorized as a product of several conditional distributions denotes the dependency/independency structure by a DAG:

$$P(X_1, \dots, X_n) = \prod_{i=1}^n P(X_i | Pa(X_i^G)). \quad (3)$$

Equation (3) (where $Pa(X_i^G)$ denotes the parent nodes of X_i) is the main reason for the formulation of a multivariate distribution by BNs; this equation is also called the *chain rule for Bayesian networks*.

In order to obtain the DAG, we used the *bnlearn* package [43] of the R language [44]. To obtain the structure, two options either select a single *best* model or obtain some *average* model, which is known as *model averaging* [45]. Our model was learned by *tabu* algorithm, which explore the search space starting from a network structure and adding, deleting or reversing one arc at a time until the score can no longer be improved. The final model was obtained by repeating several times the structure learning; a large number of network structures was explored (1000 BNs) to reduce the impact of locally optimal (but globally suboptimal) network learning. The networks learned were averaged to obtain a more robust model. The averaged network structure was obtained using the arcs present in at least 85% of the networks, which gives a measure of the strength of each arc and establishes its *significance* given a *threshold* (85%) (see Figure 1 and Listing 1).

Parameters were obtained again with the *bnlearn* package in the R language by performing a Bayesian parameter estimation using a Dirichlet distribution [46]. A conditional probability distribution is obtained for each node.

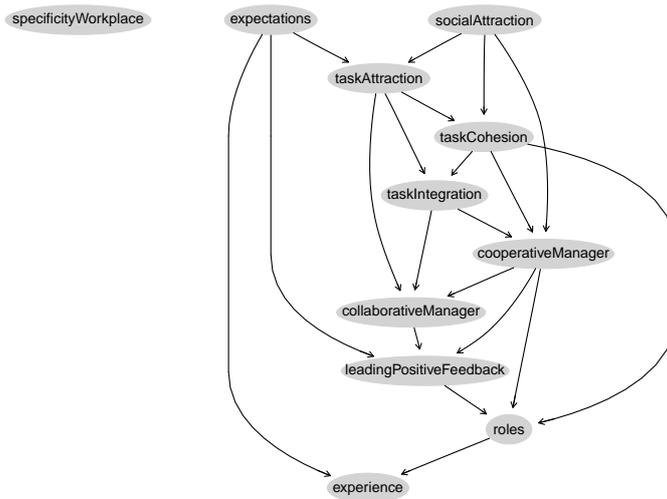


Figure 1. Structure obtained by *model averaging* over 1000 networks. It was built with the tabu learning algorithm from the *bnlearn* package in the R language using a threshold = 0.85.

2.3.3. Naive Bayes

The NB model assumes that instances fall into one of a number of mutually exclusive classes, and it is the simplest BN classifier, where the predictive variables are assumed to be conditionally independent given the class:

$$P(\mathbf{x}_i | c_j) = P(x_{i1} | c_j) \cdots P(x_{in} | c_j). \tag{4}$$

Even though the assumption of conditional independence is violated on numerous occasions in real applications, NB still performs well in many situations [41]. In Figure 2, three NB structures for *leading positive feedback*, *cooperative manager* and *collaborative manager* are obtained from the *bnlearn* package [43] in the R language [44] (see also Listing 2).

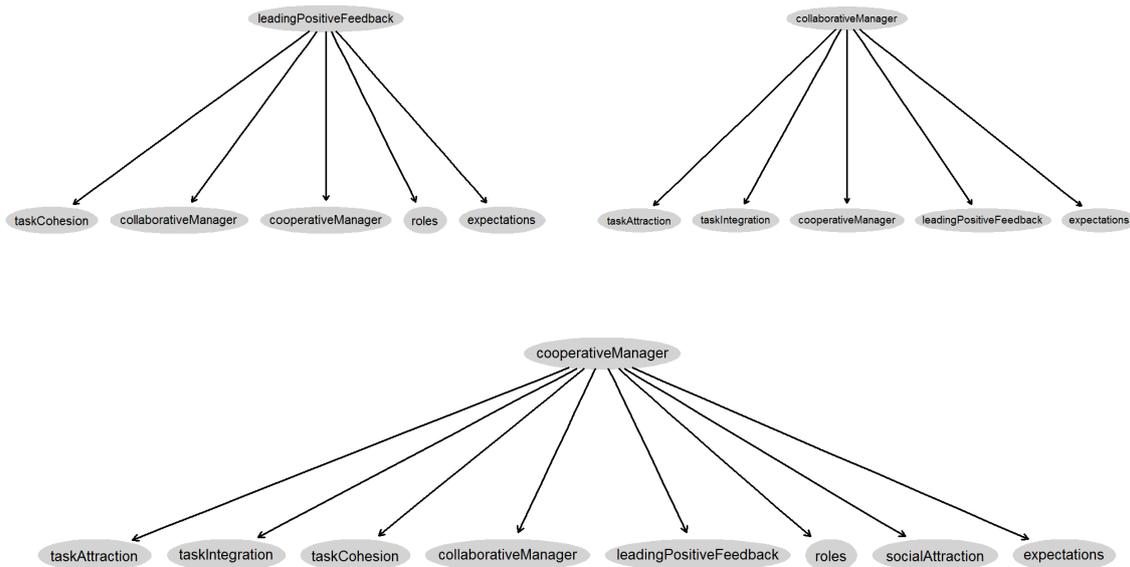


Figure 2. NB structures obtained using the *bnlearn* package in the R language for *leading positive feedback*, *cooperative manager* and *collaborative manager*.

The joint probability distribution for the NB models in Figure 2 by simple probabilistic reasoning and then the conditional independence assumption for a NB model are obtained as follows.

Suppose $C = leadingPositiveFeedback$, $X_1 = expectations$, $X_2 = collaborativeManager$, $X_3 = cooperativeManager$, $X_4 = roles$, $X_5 = taskCohesion$, then:

$$P(C, X_1, X_2, X_3, X_4, X_5) = P(X_1, X_2, X_3, X_4, X_5|C) \cdot P(C) = \prod_{i=1}^5 P(X_i|C) \cdot P(C). \tag{5}$$

If we suppose $C = cooperativeManager$, $X_1 = taskCohesion$, $X_2 = socialAttraction$, $X_3 = taskIntegration$, $X_4 = collaborativeManager$, $X_5 = leadingPositiveFeedback$, $X_6 = roles$, $X_7 = expectations$, $X_8 = taskAttraction$, then we can have:

$$P(C, X_1, X_2, X_3, X_4, X_5, X_6, X_7, X_8) = P(X_1, X_2, X_3, X_4, X_5, X_6, X_7, X_8|C) \cdot P(C) = \prod_{i=1}^8 P(X_i|C) \cdot P(C). \tag{6}$$

Finally, we can assume that $C = collaborativeManager$, $X_1 = taskAttraction$, $X_2 = taskIntegration$, $X_3 = cooperativeManager$, $X_4 = leadingPositiveFeedback$, $X_5 = expectations$; we can obtain that:

$$P(C, X_1, X_2, X_3, X_4, X_5) = P(X_1, X_2, X_3, X_4, X_5|C) \cdot P(C) = \prod_{i=1}^5 P(X_i|C) \cdot P(C). \tag{7}$$

After the above steps, we have factorized the joint distribution as a product of conditional probability distributions. In an NB model, given the parent node, the remaining nodes are independent. Given the parent node, for each pair of nodes X_i and X_j , by applying the symmetry property, the X_i node is independent of X_j and reciprocally.

2.3.4. Tree Augmented Naive Bayes

The TAN classifier [38] extends the NB model with a tree-shaped graph across the predictor variables. The TAN model is similar to NB, except that each predictor variable is allowed to depend on other predictor variables in addition to the class. Figure 3 shows three models that we have learned for *leading positive feedback*, *cooperative manager* and *collaborative manager* (see also Listing 3). This model provides more information than the NB model, as it included information about the relationship among all predictor variables.

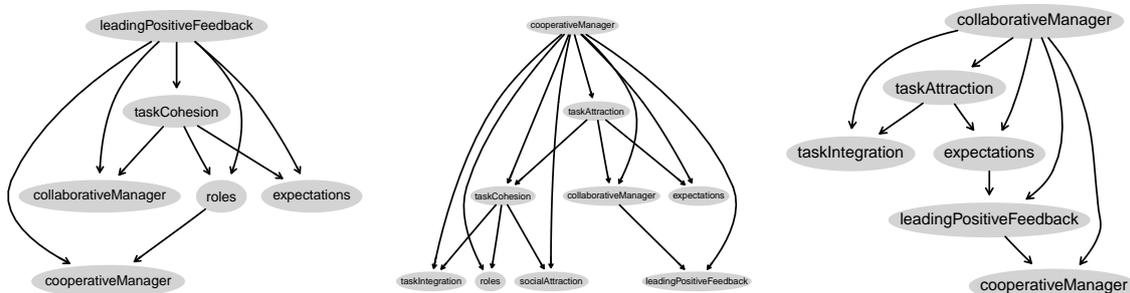


Figure 3. Three tree augmented naive Bayes (TAN) structures obtained using the *bnlearn* package in the R language for *leading positive feedback*, *cooperative manager* and *collaborative manager*.

Let us note that the joint probability distribution for the TAN models in Figure 3 by simple probabilistic reasoning and then using the conditional independence assumption for a TAN model is as follows (from left to right):

We can assume that $C = \text{leadingPositiveFeedback}$, $X_1 = \text{expectations}$, $X_2 = \text{collaborativeManager}$, $X_3 = \text{cooperativeManager}$, $X_4 = \text{roles}$, $X_5 = \text{taskCohesion}$:

$$P(X_1, X_2, X_3, X_4, X_5) = P(X_5|C) \cdot P(X_2|X_5, C) \cdot P(X_4|X_5, C) \cdot P(X_1|X_5, C) \cdot P(X_3|X_4, C) \cdot P(C). \quad (8)$$

Now, let us suppose $C = \text{cooperativeManager}$, $X_1 = \text{taskCohesion}$, $X_2 = \text{socialAttraction}$, $X_3 = \text{taskIntegration}$, $X_4 = \text{collaborativeManager}$, $X_5 = \text{leadingPositiveFeedback}$, $X_6 = \text{roles}$, $X_7 = \text{expectations}$, $X_8 = \text{taskAttraction}$, then:

$$P(C, X_1, X_2, X_3, X_4, X_5, X_6, X_7, X_8) = P(X_8|C) \cdot P(X_1|C, X_8) \cdot P(X_4|C, X_8) \cdot P(X_7|C, X_8) \cdot P(X_3|C, X_1) \cdot P(X_6|C, X_1) \cdot P(X_2|C, X_1) \cdot P(X_5|C, X_4) \cdot P(X_1). \quad (9)$$

As another step, we can assume that $C = \text{collaborativeManager}$, $X_1 = \text{taskAttraction}$, $X_2 = \text{taskIntegration}$, $X_3 = \text{cooperativeManager}$, $X_4 = \text{leadingPositiveFeedback}$, $X_5 = \text{expectations}$; then, we will obtain:

$$P(C, X_1, X_2, X_3, X_4, X_5) = P(X_1|C) \cdot P(X_2|X_1, C) \cdot P(X_5|X_1, C) \cdot P(X_4|X_5, C) \cdot P(X_3|X_4, C) \cdot P(C). \quad (10)$$

Up to this point, we obtain the probabilistic models factorized according to the TAN models, which are special types of BNs; see Figure 3. The independence of the *social attraction* and *leading positive feedback* features given the *roles* and *cooperative manager* features is shown in the center of Figure 3; reciprocally by symmetry, *leading positive feedback* is independent of the *social attraction* given the *roles* and *cooperative manager* features. Similarly, on the right side of Figure 3, *cooperative manager* and *task integration* appear as independent features given the *collaborative manager* and *task cohesion* features.

2.3.5. Validation

The models are validated using a 10-fold cross-validation. In Table 1, the area under the ROC curve (AUC) and the percentage correctly classified for the different features in the BN model are shown.

Table 1. AUCs and the percentage correctly classified for the different features by the Bayesian network (BN).

Variable Name	AUC	Accuracy
Task Attraction	0.8834 (± 0.0038)	87.3450% (± 0.2280)
Task Integration	0.8898 (± 0.0034)	90.5707% (± 0.4290)
Task Cohesion	0.9747 (± 0.0033)	93.3002% (± 0.3833)
Collaborative Manager	0.8011 (± 0.0023)	87.0968% (± 0.3133)
Cooperative Manager	0.7745 (± 0.0041)	84.6154% (± 0.3255)
Leading Positive Feedback	0.7851 (± 0.0026)	90.0744% (± 0.3265)
Experience	0.6250 (± 0.0110)	80.6951% (± 0.3519)
Roles	0.7450 (± 0.0037)	95.7816% (± 0.3050)
Social Attraction	0.6772 (± 0.0103)	78.4119% (± 0.3400)
Expectations	0.7100 (± 0.0030)	76.4268% (± 0.3781)

Standard deviation between brackets.

2.3.6. Performance Comparison

In order to provide reference benchmarks about how our BN classifies the features, we compare it to other classification performances. The performance of each classification model is evaluated using four statistical measures: accuracy, sensitivity, specificity and precision. Classification accuracy

is defined as the ratio of the number of correctly-classified cases. The sensitivity refers to the rate of correctly classified as positive. The specificity refers to the rate of correctly classified as negative, and the precision refers to the rate of positive predicted value. Other classifiers have also been considered to compare to the Bayesian classifiers: the multilayer perceptron (MP) [48], logistic regression (LR) [49], Id3 [50] and random forest (RF) [51] algorithms. The best performance is obtained by the average BN and TAN models (see Tables 2–4). However, some of these models also show a low specificity for the collaborative manager model and leading positive feedback by increasing the Type I error (one-specificity).

Table 2. Comparison of the accuracy and other metrics for the *leading positive feedback* feature using different algorithms. MP, multilayer perceptron.

Algorithms	Accuracy	Sensitivity	Specificity	Precision
BN	90.07% (± 0.3265)	0.9178 (± 0.0024)	0.6538 (± 0.0039)	0.9746 (± 0.0032)
NB	88.09% (± 0.3519)	0.8810 (± 0.0035)	0.6470 (± 0.0005)	0.8590 (± 0.0033)
TAN	90.07% (± 0.3342)	0.9156 (± 0.0030)	0.6667 (± 0.0053)	0.9775 (± 0.0048)
MP	88.83% (± 0.3267)	0.8880 (± 0.0033)	0.6640 (± 0.0232)	0.8660 (± 0.0059)
LR	88.09% (± 0.3781)	0.8810 (± 0.0038)	0.7550 (± 0.0269)	0.8480 (± 0.0082)
Id3	88.09% (± 0.3708)	0.8810 (± 0.0038)	0.7550 (± 0.0247)	0.8480 (± 0.0132)
RF	88.09% (± 0.3400)	0.8810 (± 0.0034)	0.7550 (± 0.0356)	0.8480 (± 0.0073)

Standard deviation between brackets.

Table 3. Comparison of the accuracy and other metrics for the *cooperative manager* feature using different algorithms.

Algorithms	Accuracy	Sensitivity	Specificity	Precision
BN	82.63% (± 0.3255)	0.8329 (± 0.0036)	0.7500 (± 0.0062)	0.9748 (± 0.0058)
NB	75.68% (± 0.4598)	0.7570 (± 0.0047)	0.5400 (± 0.0054)	0.7420 (± 0.0041)
TAN	83.87% (± 0.3638)	0.8462 (± 0.0043)	0.7692 (± 0.0028)	0.9716 (± 0.0022)
MP	76.67% (± 1.1174)	0.7670 (± 0.0110)	0.5460 (± 0.0194)	0.7470 (± 0.0103)
LR	79.40% (± 0.6455)	0.7940 (± 0.0064)	0.5730 (± 0.0114)	0.7650 (± 0.0088)
Id3	80.89% (± 0.6129)	0.8090 (± 0.0061)	0.5690 (± 0.0171)	0.7840 (± 0.0094)
RF	80.40% (± 1.1219)	0.8040 (± 0.0113)	0.5280 (± 0.0203)	0.7820 (± 0.0138)

Standard deviation between brackets.

Table 4. Comparison of the accuracy and other metrics for the *collaborative manager* feature using different algorithms.

Algorithms	Accuracy	Sensitivity	Specificity	Precision
BN	87.10% (± 0.3642)	0.9829 (± 0.0068)	0.5000 (± 0.0108)	0.9829 (± 0.0062)
NB	84.86% (± 0.2280)	0.8490 (± 0.0023)	0.7100 (± 0.0139)	0.8180 (± 0.0036)
TAN	87.09% (± 0.3826)	0.8824 (± 0.0036)	0.5000 (± 0.0042)	0.9829 (± 0.0096)
MP	86.85% (± 0.4290)	0.8680 (± 0.0041)	0.8060 (± 0.0233)	0.8220 (± 0.0101)
LR	86.85% (± 0.3813)	0.8680 (± 0.0033)	0.8220 (± 0.0007)	0.8180 (± 0.0081)
Id3	86.85% (± 0.3139)	0.8680 (± 0.0030)	0.8390 (± 0.0119)	0.8130 (± 0.0118)
RF	86.35% (± 0.5290)	0.8640 (± 0.0053)	0.7570 (± 0.0247)	0.8230 (± 0.0094)

Standard deviation between brackets.

Furthermore, BNs are able to produce probability estimates. In this sense, we are interested in knowing the features with the highest influence in maximizing *collective efficacy*, *cooperative manager* and *collaborative manager* in the *high* and *low* states.

2.3.7. Conditional Entropy

In Shannon [47] theory, the entropy of X is the lower bound on the average number of bits that are needed to encode values of X . Another way of viewing the entropy is as a measure of our uncertainty about the value of X , *i.e.*, lower uncertainty about X will produce a low entropy value.

A natural question is what is the cost of encoding X if we have already encoded Y . The conditional entropy of X given Y is:

$$H_P(X|Y) = E_P \left[\log \frac{1}{P(X|Y)} \right] = \sum P(X|Y) \cdot \log \frac{1}{P(X|Y)} \quad (11)$$

which captures the additional cost (in terms of bits) of encoding X when we have already encoded Y . Note that the maximum value of probability in $P(X|Y)$ implies the lowest entropy value.

For the *leading positive feedback*, *cooperative manager* and *collaborative manager* features, we are interested in determining and ordering the state values for conditioned features, because we can obtain the maximum probability value in the *low* state and the *high* state. This will lead to achieving the desired minimum conditioned entropy.

3. Results

The BN model has been selected to rank the features that can mostly increase the probability value of *leading positive feedback*, or *cooperative manager*, or *collaborative manager* in a specific state value.

Given the evidence $\mathbf{E} = e$, our goal is to find the most likely assignment to the remaining variables, denoted by \mathbf{U} ; see Equation (12):

$$MAP(\mathbf{U} | e) = \operatorname{argmax}_u P(u, e) \quad (12)$$

3.1. BN Model for Leading Positive Feedback

We select at each step the feature that maximize the most *leading positive feedback* feature likelihood in a *high* state, *i.e.*, we choose from each step the variable and the state that induces the greatest increase in the likelihood of *leading positive feedback* variable in a *high* state. A summary is shown in Table 5 and Figure 4 (on the right side).

As shown in Table 5, we can use a step by step instantiation method to maximize the likelihood of the “leading positive feedback” feature. At Step 1, with no values set for other features in the initial BN, we can reach a likelihood of 86.7%. At the subsequent steps, we add the conditions for the instantiation of other features one by one as indicated in Table 5; in each step, the percentage of the likelihood of the “leading positive feedback” increases accordingly and finally reached 96.3%. Tables 6–10 list the same step by step method to maximize the likelihood of other features.

Table 5. Step-by-step instantiations leading to maximization of the likelihood of the leading positive feedback variable in its *high* state.

Step	Instantiated Variable	Value	Leading Positive Feedback = High
1	<i>none</i>	=	–
2	<i>cooperative manager</i>	=	High
3	<i>expectations</i>	=	Low
4	<i>collaborative manager</i>	=	High
5	<i>roles</i>	=	High
6	<i>task cohesion</i>	=	Low

Again, we choose from each step the variable and the state that induces the greatest increase in the likelihood of the *leading positive feedback* variable in a *low* state. A summary is shown in Table 6 and Figure 4 (on the left side).

Table 6. Step-by-step instantiations leading to the maximization of the likelihood of the leading positive feedback variable in its *low* state.

Step	Instantiated Variable	Value	Leading Positive Feedback = Low
1	<i>none</i>	= -	13.3%
2	<i>cooperative manager</i>	= Low	34.8%
3	<i>collaborative manager</i>	= Low	63.1%
4	<i>expectations</i>	= High	66.7%
5	<i>roles</i>	= High	68.0%
6	<i>task cohesion</i>	= Low	77.6%

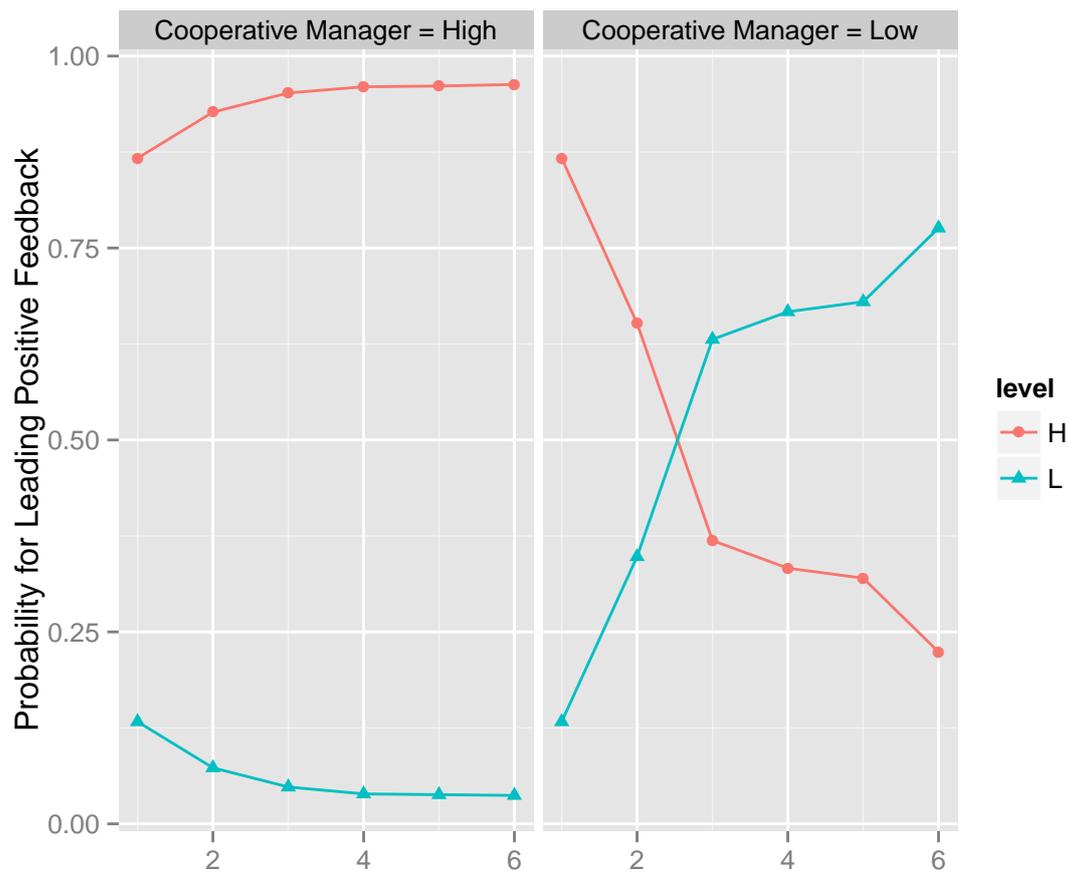


Figure 4. Step by step instantiations in the BN to maximize the *leading positive feedback* feature values in *high* and *low* states. The horizontal line represents the different steps from Tables 6 and 7.

3.2. BN Model for the Cooperative Manager

We select at each step the feature that maximize the greatest *cooperative manager* feature likelihood in a *high* state, *i.e.*, we choose from each step the variable and the state that induce the greatest increase in the likelihood of the *cooperative manager* variable in a *high* state. A summary is shown in Table 7 and Figure 5.

Table 7. Step-by-step instantiations leading to the maximization of the likelihood of the cooperative manager variable in its *high* state.

Step	Instantiated Variable	Value	Cooperative Manager = High
1	<i>none</i>	= -	78.1%
2	<i>leading positive feedback</i>	= High	83.6%
3	<i>task cohesion</i>	= High	86.0%
4	<i>social attraction</i>	= High	87.4%
5	<i>expectations</i>	= High	87.8%
6	<i>collaborative manager</i>	= High	88.5%
7	<i>task attraction</i>	= High	89.0%
8	<i>task integration</i>	= High	89.2%
9	<i>roles</i>	= High	89.3%

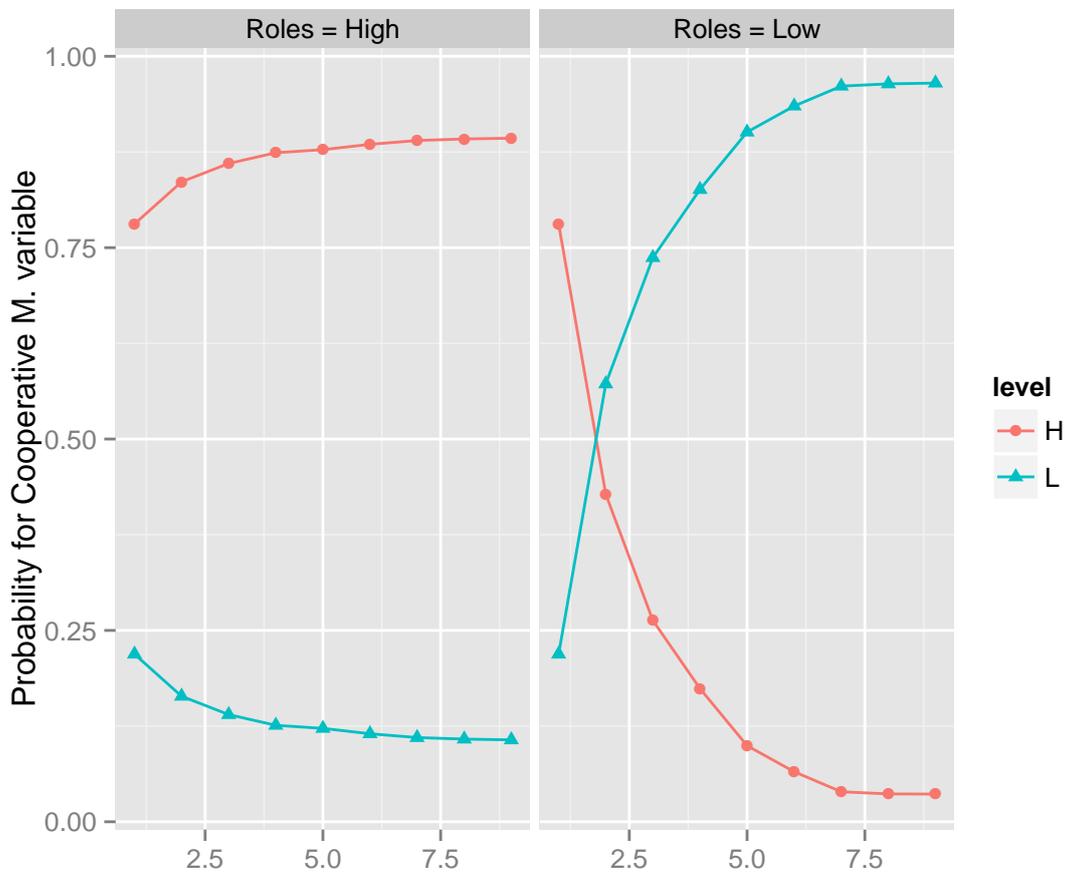


Figure 5. Step by step instantiations in the BN to maximize the *cooperative manager* feature in *high* and *low* states. The horizontal line represents the different steps from Tables 8 and 9.

Again, we choose from each step the variable and the state that induces the greatest increase in the likelihood of the *cooperative manager* variable in a *low* state. In this case, the cooperative manager achieves the maximum in the *low* state when leading positive feedback and roles are instantiated in a *low* value; however, to see the influence of the other features, we show the roles feature in the last position. A summary is shown in Table 8 and Figure 5 (on the right side).

Table 8. Step-by-step instantiations leading to the maximization of the likelihood of the cooperative manager variable in its *low* state.

Step	Instantiated Variable	Value	Cooperative Manager = Low
1	<i>none</i>	= -	21.9%
2	<i>leading positive feedback</i>	= Low	57.2%
3	<i>collaborative manager</i>	= Low	73.7%
4	<i>roles</i>	= Low	82.6%
5	<i>task attraction</i>	= High	90.1%
6	<i>task integration</i>	= High	93.5%
7	<i>social attraction</i>	= Low	96.1%
8	<i>expectations</i>	= Low	96.4%
9	<i>task cohesion</i>	= High	96.5%

3.3. BN Model for the Collaborative Manager

We select at each step the feature that maximize the greatest *collaborative manager* feature likelihood in a *high* state, *i.e.*, we choose from each step the variable and the state that induce the greatest increase in the likelihood of the *collaborative manager* variable in a *high* state. A summary is shown in Table 9 and Figure 6 (on the right side).

Table 9. Step-by-step instantiations leading to the maximization of the likelihood of the collaborative manager variable in its *high* state.

Step	Instantiated Variable	Value	Collaborative Manager = High
1	<i>none</i>	= -	85.5%
2	<i>task attraction</i>	= High	90.7%
3	<i>cooperative manager</i>	= High	93.3%
4	<i>task integration</i>	= High	95.4%
5	<i>leading positive feedback</i>	= High	95.7%
6	<i>expectations</i>	= Low	95.8%

Again, we choose from each step the variable and the state that induce the greatest increase in the likelihood of the *collaborative manager* variable in a *low* state. In this case, the cooperative manager achieves the maximum in the *low* state when the features task attraction, leading positive feedback, task cohesion, experience, cooperative manager and task integration are all instantiated in a *low* state. A summary is shown in Table 10 and Figure 6 (on the left side).

Table 10. Step-by-step instantiations leading to the maximization of the likelihood of the collaborative variable in its *low* state.

Step	Instantiated Variable	Value	Collaborative Manager = Low
1	<i>none</i>	= -	14.5%
2	<i>leading positive feedback</i>	= Low	34.9%
3	<i>task integration</i>	= Low	58.0%
4	<i>task attraction</i>	= Low	66.1%
5	<i>expectations</i>	= Low	76.3%
6	<i>cooperative manager</i>	= Low	81.4%

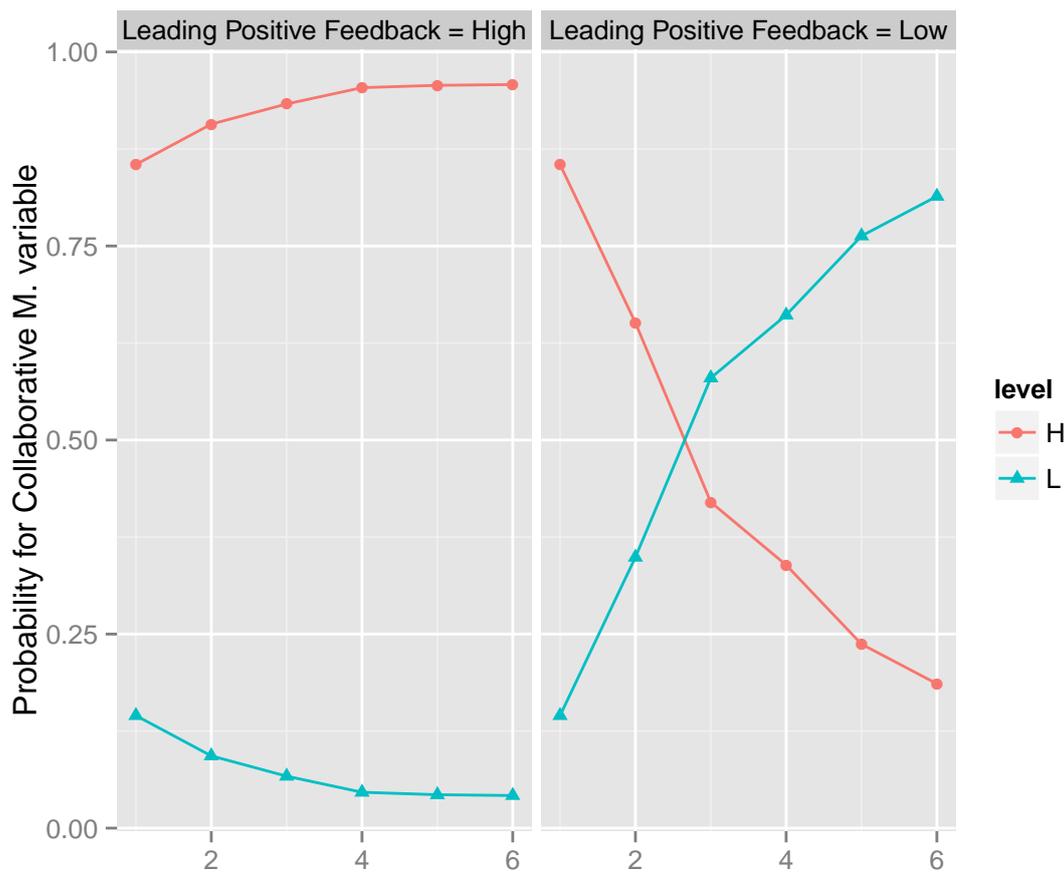


Figure 6. Step by step instantiations in the BN to maximize the *collaborative manager* feature in *high* and *low* states. The horizontal line represents the different steps from Tables 8 and 9.

4. Discussion

In this study, ranking the psychological team features is pretty relevant attending to the manager's and coaches' great difficulties in generating specific climates addressed to obtain better performance from the team members acting as a whole team [53,55].

At present, there is still discussion about the extrapolation of the mechanisms of the teams and their tasks; however, there is a common agreement on the importance of the concept of cohesion, leadership and the specific working position [9,10,54–56]. From this point of view, this set of concepts has been considered as a prototype for racing sports teams and should be discussed in terms of the relationship with the group performance. In previous works [52], the relationships between some features were established. In this study, we used the Markov blanket in the average BN to determine a different subset of features. After, two BN classifiers are determined. The aim of this study has been reached through the elaboration of the BN, the subsequent TANs and some major instantiations. Our study shows the importance of the role played by social attraction (as was predicted by Carron and Eys ([2,9,36]) and the team members' expectations, as was outlined when studying the role of self-efficacy related to the teams' performance [8]. These two variables appear as the BN antecessors, influencing the rest of the variables. Furthermore, in our study, the main node has been found to be the cooperative teamwork style (as has been outlined in previous studies [13,20,52]); it determines the BN bottom consequences, the members roles and their experience, *i.e.*, their time duration for playing on the same team. These two consequences are critical for the determination of the team performance and their stability across different situations. However, there may be a distortion of objectivity due to the team affective attraction, which can possibly affect the members' trust of the

team's capabilities [53]. When we conduct instantiations on the BN and maximize leading positive feedback managerial behavior in *high* and *low* states, BN showed the cooperative manager as the one with the strongest influence. A very interesting finding was that the initial expectations about the team performance need to be *low* in case the leading positive feedback is maximized to *high*, and it needs to be *high* in case the leading positive feedback is maximized to *low*, in agreement with previous work [20]. After that, knowing under which features the cooperative manager would work with maximum probability in *high* and *low* states is quite important. When we instantiated the cooperative manager feature in the maximum of *high* and *low* states, BN showed, not surprisingly, leading positive feedback as the one with the strongest influence. Adding roles in *high* and *low* states, we maximize the cooperative manager in *high* and *low* states, respectively, achieving an entropy with a value of zero. The other variables that have shown influence on the cooperative managerial style are the social attraction, the specificity of the workplace and the roles when the style is maximized to both *low* and *high* states. Maximizing the collaborative manager feature in *high* and *low* states is also of special interest, showing its contribution to the main cooperative node. It appears that when we maximize the collaborative manager in a *high* state, the task attraction is the variable with the strongest influence, followed by the cooperative manager, task integration (instantiated in the *high* state), leading positive feedback and, finally, expectations (when team members behave as being focused just on reaching the team's objectives). On the other hand, the task integration can be instantiated in a *low* state in case we are interested in maximizing the collaborative manager to the *low* state.

5. Conclusions

In the study of sports science and other cooperative work systems, BNs provide some major advantages. They explicitly provide the conditional probability distributions of the values of every feature given the values of other input features. They are represented in a clear, appealing way, and are very easy to comprehend and translate to end users, *i.e.*, managers and coaches. In this study, by the ranking obtained from the BN among the psychological variables, we come to the conclusion that the cooperative style with its associated behaviors constitutes the most prominent psychological variable. Managers have to work with team members' expectations and the social aspects of the job, because they are the antecessors of the whole BN, and also, they need to keep in mind that the cooperation, developed mainly through a positive feedback work climate and a pro-social, *i.e.*, collaborative, managerial style, leads to the correct decision of the proper assignation of team roles and the working duration of the team members.

Acknowledgments: This study was supported by a project of the Spanish Ministry of Economy and Competitiveness (TIN2013-42795-P).

Author Contributions: Conceived of and designed the experiments: A.G.-M., P.F.P., F.J.P., J.C., Y.L. performed the experiments: P.F.P., A.G.-M., F.J.P., J.C. Analyzed the data: P.F.P. Contributed reagents/materials/analysis tools: P.F.P., A.G.-M., Y.L. Wrote the manuscript: P.F.P., A.G.-M., F.J.P., Y.L. All authors contributed to critically revising the submitted version.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Dataset

The dataset is available online as supporting material.

Appendix B. The R code

Listing 1: R code

```
# Load libraries
library(Rgraphviz)
library(bnlearn)

#Download the data
soccerR <- read.table("D:/dataset01.txt", header=T, quote="\")

#Average Bayesian network
start = random.graph(nodes=names(soccerR), num=1000)
netlist = lapply(start, function(net) {
  tabu(soccerR, score = "aic", start = net)
})
arcs = custom.strength(netlist, nodes = names(soccerR), cpdag = FALSE)
arcs[(arcs$strength > 0.85) & (arcs$direction >= 0.5),]
modelstring(averaged.network(arcs))
plot(averaged.network(arcs))

#Plot with Graphviz
graphviz.plot(averaged.network(arcs), highlight=list(nodes=c("cooperativeManager",
"specificityWorkplace", "experience", "taskAttraction", "taskIntegration",
"socialAttraction", "taskCohesion", "collaborativeManager", "leadingPositiveFeedback",
"roles", "expectations"), col=c("lightgrey"), fill=c("lightgrey")), layout="dot",
shape="ellipse", main=NULL, sub=NULL)
```

Listing 2: R code

```
bncoll = naive.bayes(coll, "collaborativeManager")
bncoop = naive.bayes(coop, "cooperativeManager")
bnlead = naive.bayes(lead, "leadingPositiveFeedback")
```

Listing 3: R code

```
tancoll = tree.bayes(coll, "collaborativeManager")
tancoop = tree.bayes(coop, "cooperativeManager")
tanlead = tree.bayes(lead, "leadingPositiveFeedback")
```

References

1. Callow, N.; Smith, M.J.; Hardy, L.; Arthur, C.A.; Hardy, J. Measurement of transformational leadership and its relationship with team cohesion and performance level. *J. Appl. Sport Psychol.* **2009**, *22*, 395–412.
2. Carron, A.V.; Brawley, L.R.; Widmeyer, W.N. Measurement of cohesion in sport and exercise. In *Advances in Sport and Exercise Psychology Measurement*; Duda, J.L., Ed.; Fitness Information Technology: Morgantown, WV, USA, 1998; pp. 213–226.
3. Heuzé, J.P.; Raimbault, N.; Fontayne, P. Relationships between cohesion, collective efficacy and performance in professional basketball teams: An examination of mediating effects. *J. Sports Sci.* **2006**, *24*, 59–68.
4. Myers, N.D.; Payment, C.A.; Feltz, D.L. Reciprocal relationships between collective efficacy and team performance in women's ice hockey. *Group Dyn.-Theor. Res.* **2004**, *8*, 182–195.
5. Wilson, D.S.; Ostrom, E.; Cox, M.E. Generalizing the core design principles for the efficacy of groups. *J. Econ. Behav. Organ.* **2013**, *90S*, S21–S32.
6. Mathieu, J.; Maynard, M.T.; Rapp, T.; Gilson, L. Team effectiveness 1997–2007: A review of recent advancements and a glimpse into the future. *J. Manag.* **2008**, *34*, 410–476.
7. Cannon-Bowers, J.A.; Salas, E. Teamwork competencies: The interaction of team members knowledge, skills, and attitudes. In *Workforce Readiness: Competencies and Assessment*; O'Neil, H.F., Ed.; Erlbaum: Mahwah, NJ, USA, 1997; pp. 151–174.
8. Fuster-Parra, P.; García-Mas, A.; Ponseti, F.J.; Leo, F.M. Team performance and collective efficacy in the dynamic psychology of competitive team: A Bayesian network analysis. *Hum. Mov. Sci.* **2015**, *40*, 98–118, doi:10.1016/j.humov.2014.12.005.
9. Beal, D.J.; Cohen, R.R.; Burke, M.J.; McLendon, C.L. Cohesion and performance in groups: A meta-analytic clarification of construct relations. *J. Appl. Psychol.* **2003**, *88*, 989–1004.
10. Carron, A.V.; Bray, S.R.; Eys, M.A. Team cohesion and team success in sport. *J. Sports Sci.* **2002**, *20*, 119–126.

11. Eys, M.A.; Carron, A.V.; Beauchamp, M.R.; Bray, S.R. Role ambiguity in sport teams. *J. Sport Exerc. Psychol.* **2003**, *25*, 534–550.
12. Jirotko, M.; Gilbert, N.; Luff, P. On the social organisation of organisations. *Comput. Support. Coop. Work J.* **1992**, *1*, 95–118.
13. Lameiras, J.; Almeida, P.L.; García-Mas, A. Relationships between cooperation and goal orientation among male professional and semi-professional team athletes. *Percept. Motor Skills* **2014**, *119*, 851–860.
14. Heckerman, D. Bayesian networks for data mining. *Data Min. Knowl. Discov.* **1997**, *1*, 79–119.
15. Jensen, F.V.; Nielsen, T.D. *Bayesian Networks and Decision Graphs*; Information Science & Statistics; Springer: Berlin/Heidelberg, Germany; New York, NY, USA, 2007.
16. Koller, D.; Friedman, N. *Probabilistic Graphical Models. Principles and Techniques*; The MIT Press: Cambridge, MA, USA; London, UK, 2010.
17. Korb, K.B.; Nicholson, A.E. *Bayesian Artificial Intelligence*; Chapman and Hall/CRC Press: London, UK, 2010.
18. Pearl, J. *Causality. Models, Reasoning and Inference*; Cambridge University Press: Cambridge, UK, 2000.
19. Fuster-Parra, P.; García-Mas, A.; Ponseti, F.J.; Palou, P.; Cruz, J. A Bayesian network to discover relationships between negative features in sport: A case study of teen players. *Qual. Quant.* **2014**, *48*, 1473–1491, doi:10.1007/s11135-013-9848-y.
20. Fuster-Parra, P.; García-Mas, A.; Cantalops, J.; Ponseti, F.J. Cooperative team work analysis and modeling: A Bayesian network approach. In *Cooperative Design, Visualization and Engineering*; Luo, Y., Ed.; Springer International Publishing: Berlin, Germany, 2015; pp. 1–10, doi:10.1007/978-3-319-24132-6-1.
21. Fuster-Parra, P.; Tauler, T.; Bennasar-Veny, M.; Ligeza, A.; López-González, A.A.; Aguiló, A. Bayesian network modeling: A case study of an epidemiologic system analysis of cardiovascular risk. *Comput. Methods Progr. Biomed.* **2016**, *126*, 128–142.
22. DeFelipe, J.; López-Cruz, P.L.; Benavides-Piccione, R.; Bielza, C.; Larrañaga, P.; Anderson, S.; Burkhalter, A.; Cauli, B.; Fairén, A.; Feldmeyer, D.; et al. New insights into the classification and nomenclature of cortical GABAergic interneurons. *Nat. Neurosci.* **2013**, *14*, 202–216.
23. Ligeza, A.; Fuster-Parra, P. AND/OR/NOT causal graphs. A model for diagnostic reasoning. *Int. J. Appl. Math. Comput. Sci.* **1997**, *7*, 185–203.
24. Cooper, G.F.; Herskovits, E. A Bayesian method for the induction of probabilistic networks from data. *Mach. Learn.* **1992**, *9*, 309–347.
25. Heckerman, D.; Geiger, D.; Chickering, D.M. Learning Bayesian networks: The combination of knowledge and statistical data. *Mach. Learn.* **1995**, *20*, 197–243.
26. Butz, C.J.; Hua, S.; Chen, J.; Yao, H. A simple graphical approach for understanding probabilistic inference in Bayesian networks. *Inf. Sci.* **2009**, *179*, 699–716.
27. Glymour, C. *The Mind's Arrows: Bayes Nets and Graphical Causal Models in Psychology*; The MIT Press: New York, NY, USA, 2003.
28. Druzdzel, M.J.; Glymour, C. What do college ranking data tell us about student retention: Causal discovery in action. Intelligent Information Systems IV. In Proceedings of the Workshop Held in Augustów, Poland, 5–9 June 1985; pp. 1–10.
29. Glymour, C.; Scheines, R.; Spirtes, P.; Kelly, K. *Discovering Causal Structure*; Academic Press: New York, NY, USA, 1987.
30. Spirtes, P.; Glymour, C.; Scheines, R. *Causation, Prediction and Search*, 2nd ed.; Adaptive Computation and Machine Learning; The MIT Press: New York, NY, USA, 2001.
31. Darwiche, A. Bayesian networks. In *Handbook of Knowledge Representation, Foundations of Artificial Intelligence*; Van Harmelen, A., Lifschitz, V., Porter, B., Eds.; Elsevier: Amsterdam, The Netherlands, 2008; p. 1034.
32. Witten, I.H.; Frank, E. *Data Mining: Practical Machine Learning Tools and Techniques*; Morgan Kaufman: Amsterdam, The Netherlands, 2005.
33. Felipe, S.C.; Pires, D.S.; Nassar, S.M. Analysis of bayesian classifier accuracy. *J. Comput. Sci.* **2013**, *9*, 1487–1495.
34. García Mas, A.; Fuster-Parra, P.; Ponseti, J.; Palou, P.; Olmedilla, A.; Cruz, J. Análisis de las relaciones entre motivación, el clima motivacional y la ansiedad competitiva entre jóvenes jugadores de equipo mediante una red Bayesiana. *An. Psicol.-Spain* **2015**, *1*, 355–366, doi:10.6018/analesps.31.1.167531.
35. Crespo, M.; Balaguer, I.; Atienza, F.L. Análisis psicométrico de la versión española de la escala de liderazgo en el deporte de Chelladurai y Saleh en la versión de entrenadores. *Rev. Psicol. Soc. AP* **1994**, *4*, 5–28.

36. Carron, A.V.; Eys, M.A. *Group Dynamics in Sport*; Fitness Information Technology: Morgantown, WV, USA, 2012.
37. Iturbide, L.M.; Elosua, P.; Yanes, F. Medida de la cohesión en equipos deportivos. Adaptación al español del Group Environment Questionnaire (GEQ). *Psicothema* **2010**, *22*, 482–488.
38. Friedman, N.; Geiger, D.; Goldszmidt, M. Bayesian network classifiers. *Mach. Learn.* **1997**, *29*, 131–163.
39. Duda, R.; Hart, P.; Stork, D. *Pattern Classification*; John Wiley and Sons: New York, NY, USA, 2001.
40. Bielza, C.; Larrañaga, P. Discrete Bayesian network classifiers: A survey. *ACM Comput. Surv.* **2014**, *47*, 5:1–5:43.
41. Domingos, P.; Pazzani, M. On the optimality of the simple Bayesian classifier under zero-one loss. *Mach. Learn.* **1997**, *29*, 103–130.
42. Nagarajan, R.; Scutari, M.; Lèbre, S. *Bayesian Networks in R: With Applications in Systems Biology*; Springer: Berlin/Heidelberg, Germany; New York, NY, USA, 2013.
43. Scurati, M. Learning Bayesian networks with the bnlearn R package. *J. Stat. Softw.* **2010**, *35*, 1–22.
44. R Development Core Team. *R: A Language and Environment for Statistical Computing*; [Computer Software]; R Foundation for Statistical Computing: Vienna, Austria. Available online: <http://www.R-project.org/> (accessed on 20 March 2012).
45. Claeskens, G.; Hjort, N.L. *Model Selection and Model Averaging*; Cambridge University Press: Cambridge, UK, 2008.
46. Neapolitan, N.E. *Learning Bayesian Networks*; Prentice Hall, Inc.: Upper Saddle River, NJ, USA, 2003.
47. Shannon, C.E. A mathematical theory of communication. *Bell. Labs Tech. J.* **1948**, *27*, 379–423, doi:10.1002/j.1538-7305.1948.tb01338.x.
48. Haykin, S. *Neural Networks: A Comprehensive Foundation*; Prentice Hall: New York, NY, USA, 1989.
49. Hosmer, D.W.; Lemeshow, S. *Applied Logistic Regression*; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2000.
50. Quinlan, J.R. Induction of decision trees. *Mach. Learn.* **1986**, *1*, 81–106.
51. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32.
52. Olmedilla, A.; Ortega, E.; Almeida, P.; Lameiras, J.; Villalonga, T.; Sousa, C.; Torregrosa, M.; Cruz, J.; Garcia-Mas, A. Cohesión y cooperación en equipos deportivos. *An. Psicol.* **2011**, *27*, 232–238.
53. Costa, A. Work team Trust and effectiveness. *Pers. Rev.* **1994**, *32*, 605–672.
54. Cohen, S.G.; Bailey, D.E. What makes team work: Group effectiveness research from the shop floor to the executive suite. *J. Manag.* **1997**, *23*, 239–290.
55. Rico, R.; Alcover de la Hera, C.M.; Tabernero, C. Efectividad de los equipos de trabajo: Una revisión de la última década de investigación (1999–2009) (Effectiveness of Workteams: A review over the last decade of research (1999–2000)). *Rev. Psicol. Trab. Organ.* **2010**, *26*, 47–71.
56. Taggar, S.; Brown, T.C. Problem-solving team behaviors: Development and validation of BOS and a hierarchical factor structure. *Small Group Res.* **2001**, *32*, 698–726.



© 2016 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) license (<http://creativecommons.org/licenses/by/4.0/>).