Noname manuscript No.
(will be inserted by the editor)

Using a Case-Based Reasoning Approach for Trading in Sports Betting Markets

Juan M. Alberola · Ana Garcia-Fornes

Received: date / Accepted: date

Abstract The sports betting market has emerged as one of the most lucrative markets in recent years. Trading in sports betting markets entails predicting odd movements in order to bet on an outcome, whilst also betting on the opposite outcome, at different odds in order to make a profit, regardless of the final result. These markets are mainly composed by humans, which take decisions according to their past experience in these markets. However, human rational reasoning is limited when taking quick decisions, being influenced by emotional factors and offering limited calibration capabilities for estimating probabilities. In this paper, we show how artificial techniques could be applied to this field and demnostrate that they can outperform even the bevahior of high-experienced humans. To achieve this goal, we propose a case-based reasoning model for trading in sports betting markets, which is integrated in an agent to provide it with the capabilities to take trading decisions based on future odd predictions. In order to test the performance of the system, we compare trading decisions taken by the agent with trading decisions taken by human traders when they compete in real sports betting markets.

Keywords Sports Betting Markets · Trading · Odds Prediction · Case-Based Reasoning · Humans

1 Introduction

In the last few years, sports betting markets [6] have emerged as one of the scenarios in which the most money is exchanged every day. Sports betting markets are a specific kind of prediction market where the traded assets are related to sporting events. The price of these assets reflects the odds, which are related to the probability of each outcome. Therefore, the attraction of betting on sporting events and the growth in popularity that it has experienced have meant that millions of users make more exchanges in sports betting markets on an average day than in other exchange scenarios such as financial markets [8].

Prediction markets have been studied as powerful mechanisms for predicting the probabilities of future events. Most of the research on prediction markets is focused on pricing, that is, assessing the most accurate odds according the probability of the event. Studies regarding prediction markets, such as the Iowa Electronic Markets, the Foresight Exchange or the Hollywood Stock Exchange demonstrate than these markets provide very accurate probability predictions of future outcomes [15,33,9]. Other works are focused on studying how information is incorporated into the market and therefore, influences the odds [28,10]. However, to our knowledge little effort has been made towards studying odds evolution in prediction markets, with the aim of making a profit regardless of the correct outcome.

Studying how probabilities and odds, change during the sporting event will allow us to approach sports betting markets in a novel way that is similar to the financial markets. Just as financial traders buy and sell assets according to the price of assets and their expectations of price increases or decreases, assets related to sporting event can also be traded at a given odds in order to make an opposite trade later in the same market at better odds with the goal of making a profit.

These markets are mainly composed by humans with limited capabilities for prediction and reasoning. Hu-

Departament de Sistemes Informàtics i Computació
Universitat Politècnica de València
Camí de Vera s/n. 46022, València. Spain
E-mail: jalberola@dsic.upv.es

mans usually take decisions based on what they expect that is going to happen in the next minutes according what they observed in past events. They predict whether the odds will move up or down, how much it will move up or down, when it is going to move and how fast for each one of the scenarios that can appear in the match in the next minutes. However, human rational reasoning has imperfections such as emotional factors as long as poor calibration for estimating probabilities [11]. Since probabilities and therefore odds, change rapidly during a sporting event, we want to demonstrate how artificial models not only can suit well in these scenarios but also how they can become really competitive. The higher accuracy and reasoning capabilities provided by artificial models can improve the trading capabilities of the most expert human traders.

With the trading goal in mind, the main objective of this paper is to explore the novel concept of trading in sports betting markets, and to validate the suitability of artificial models in this domain. We present a model that tries to simulate the reasoning model used by humans, which, as far as we are concerned, is based on finding pattern movements due to past experience. We are interested in identifying patterns of odd movements that can be repeated in different events under similar circumstances. This model extends our work presented in [4], which is able to capture some features of a current event and to find similarities with other past events.

The underlying mechanism of our model is a Case-Based-Reasoning (CBR) [2,34,23] model, in which by means of observing past sporting events, it can predict future odd movements for an unknown sporting event and therefore, drawing a odds evolution over time. To apply this model, we provide an agent with trading capabilities called the CBR-Trader agent (CBR-Tagent). This agent is aimed to take what it considers the best trading decisions in a rational and non-emotional way.

In order to validate our model, we present a comparison between the performance of the CBR-Tagent against human traders when they are trading in real markets. Other techniques similar to some of the used for stock markets prediction [31] such as neural networks, support vector machines, quantitative matrix's, etc., could also be used, but this work is not focused on finding the best-performance technique for trading in sports betting markets.

The rest of the article is organized as follows. Section 2 presents a short introduction to sports betting markets. In Section 3 we define the trading model for sports betting markets. Section 4 details the reasoning system of the CBR-Tagent. In Section 5 we present a comparison between the performance of the CBR-Tagent and the performance of human traders. Finally, Section 6 discusses the contributions of the paper and future work.

2 Sports Betting Markets

Sports betting markets are speculative scenarios about sporting events, where participants exchange assets regarding a specific outcome of the event. For a specific event, there are several markets, each one with n possible outcomes. The odds of the exchanged assets are related to the probability of the specific outcome happening. Odds and implicit probability of an outcome are related by $odds = 1/probability$, where probability is represented from 0 to 1. Trades are made between participants at a given odds because they have different expectations about the event.

Definition 1 (back and lay). In sports betting markets, users can bet on an outcome (win if this is the final outcome) or against it (win if any of the other outcomes is the final one). Betting on a specific outcome is called *back* and betting against it is called *lay*. The bookmaker is in charge of receiving the offers of the traders. If a *back* offer and a *lay* offer are compatible in terms of odds and stake (full or partial), the bookmaker matches both offers. If a received offer is not compatible it remains waiting until a compatible offer is received, or until it is deleted by the user or because the market is closed. Therefore, the bookmaker also maintains a list of waiting offers, and continuously shows the best *back* and *lay* odds of these waiting offers in order to allow interested users make offers at these odds.

Let us suppose that *Alice* wants to bet μ units on placing a *back* bet on a selection (an individual, a team, horse, etc.). This user is betting that the selection will win. *Alice* can accept the best waiting *lay* offer (the *lay* offer which odds are the lowest ones) or can choose her own odds ρ. When *Bob* wants to place a *lay* bet on this selection (against the selection of *Alice*, that is, the individual will not win), he can also choose his own odds or accept the offer of *Alice*. If *Bob* accepts *Alice*'s offer, *Bob* is placing a *lay* bet on the selection at odds ρ. When the event is over, if *Alice* wins the bet, *Bob* has to pay $\rho - 1$ units for each unit bet on. If *Bob* wins the bet, he keeps the μ units of *Alice*:

$$\text{profit(Alice)} = \begin{cases} \mu\times(\rho\text{-}1) & \text{if Alice wins the bet} \\ -\mu & \text{if Alice loses the bet} \end{cases}$$

$$\text{profit(Bob)} = \begin{cases} -\mu\times(\rho\text{-}1) & \text{if Alice wins the bet} \\ \mu & \text{if Alice loses the bet} \end{cases}$$

3 A Trading Model for Sports Betting Markets

In sports betting markets, the probability associated to a specific outcome changes throughout the event. Furthermore, since these markets usually have a short or very short duration (a few hours, minutes or even seconds), probabilities and therefore, odds, are continuously changing throughout the sporting event. Taking into account these odd movements, we propose the study of these markets using an approach similar to the one used in financial markets. Financial traders buy and sell assets according to the asset prices and their expectations of price increases or decreases. Our approach is to trade sports assets related to a specific outcome at a given odds in order to make an opposite trade later in the same market at better odds, with the goal of making a profit regardless of what the final outcome is. From now on, we refer to this approach as sports betting markets trading, or just trading.

Definition 2 (opening and closing bets). Trading consists in two bets: the *opening* bet, which is the first bet made according to our expectations of odd movements; and the *closing* bet, which is the opposite bet in order to make a profit (or to reduce losses if the odds have not moved as we expected) whatever the final outcome is. Thus, the main aim is to detect whether the odds will move up or down, how much it will move up or down, when it is going to move and how fast. We must point that since we carried out both bets in the same market, the profit opportunities are not risk-free, since odds change may be in the opposite direction.

A trader can make a bet by risking μ units while making a *back* bet at ρ_1 odds for a specific outcome ν. As explained in previous section, if the bet is finally won, the trader wins $\mu \times (\rho_1 - 1)$ units and loses the μ units if the other outcome is the final one. This trader can make a profit if she covers all the bases by betting on the opposite outcome when odds change. In this example, the trader can bet μ on the *lay* side at ρ_2 odds. When the event ends, if the final outcome is ν, the trader will win $\mu \times (\rho_1 - 1)$ because her first bet has won; however, she loses $\mu \times (\rho_2 - 1)$ units bet on the second one. As can be observed, if the odds of the *back* bet (ρ_1) is higher than the odds of the *lay* bet (ρ_2), the resulting profit is positive:

$$\mu \times (\rho_1 - 1) - \mu \times (\rho_2 - 1) = \mu \times (\rho_1 - \rho_2)$$

Nevertheless, if the ν outcome is not the final one, the trader will not lose any units because she loses the μ units risked in the first bet but wins μ units from the second one. Thus, regardless of whether or not ν is the final outcome, the trader will not lose any units:

$$profit = \begin{cases} \mu\times(\rho_1\text{-}\rho_2) & \text{if } \nu \text{ is the final outcome} \\ 0 & \text{otherwise} \end{cases}$$

The same operations can also be done in the inverse order, that is, first make a bet on *lay* and then, at higher odds, bet on *back*. In this case, the trader bets on *lay* μ_1 units at odds ρ_1. If the odds changes to ρ_2, the trader can bet μ_2 units on *back* in order to make a positive profit if ν is not the final outcome, and a profit of 0 otherwise:

$$\mu_1 \times (\rho_1 - 1) = \mu_2 \times (\rho_2 - 1)$$

$$\mu_2 = \frac{\mu_1 \times (\rho_1 - 1)}{(\rho_2 - 1)}$$

$$profit = \begin{cases} \mu_1\text{-}\mu_2 & \text{if } \nu \text{ is not the final outcome} \\ 0 & \text{otherwise} \end{cases}$$

When the trader closes the trade, she can also split the profits between the different outcomes by betting $\mu \times (\rho_1/\rho_2)$ units on the opposite outcome. Thus, the profits for any outcome will be:

$$\begin{cases} \mu \times (\rho_1/\rho_2) - \mu & \text{if the starting bet is } back \\ \mu - \mu \times (\rho_1/\rho_2) & \text{if the starting bet is } lay \end{cases}$$

This profit will be positive if the starting bet is *back* when $\rho_1 > \rho_2$, and with will also be positive when $\rho_2 > \rho_1$ if the starting bet is *lay*. By taking into account these concepts, we can summarize the requirements for trading in sports betting markets as:

- Prediction of the probability of an event in the next few time-steps.
- Prediction of changes in odds in the next few time-steps.
- Identifying profitable trades.
- Take trading decisions.

The prediction of the probability of an event provides us with information regarding how the odds should evolve if an event occurs in the next few time-steps. In other case, how the odds should change if this event does not occur in the next few time-steps. We manage these probabilities as boolean conditions that are true or false in the next few time-steps. Then, given these predictions, we can identify profitable trades. Finally, according to profitable trades, we must select the best option for carrying out the trade.

3.1 Prediction of Probabilities

The odds associated to a specific outcome depends on the probability of this outcome. As a simple example, in a fair coin flip game, the probabilities of heads and tails are both 0.5. In this fair game, if the coin flip is repeated n times, the occurrences of both heads and tails would be $n/2$, being the fair odds 2, because, on average, we expect to lose one unit for every unit won, where $profit = units \times (odds - 1)$. For the same game, if a participant is offering us odds higher than 2 to bet on heads (for example 2.5), according to our expectation, we should accept this trade because the expected value (E) of the profit resulting from a unit bet is positive:

$$\text{E} = (\text{lost units} \times \text{lose prob.}) + (\text{won units} \times \text{win prob.}) =$$

$$= (-1 \times 0.5) + (1 \times (2.5 - 1) \times 0.5) = 0.25$$

Definition 3 (value). We say that odd have *value* if, in the long-term, we would make a profit from betting on these. In this example, the value is in the heads.

Definition 4 (condition). In the sports betting markets trading, the value of odds must be measured in terms of the probability associated to this outcome increases or decreases. These increases or decreases are caused by what happens in the sporting event in the next few time-steps. This, can be represented as *conditions* that are true or false in the next time-steps. As an example, let us imagine a tennis match in which the probabilities for a specific player $p1$ to win the match are:

$$p = \begin{cases} \rho & \text{for } p1 \text{ to win the game} \\ (1-\rho) & \text{for } p1 \text{ not to win the game} \end{cases}$$

These probabilities are going to change during the event depending on what is happening in it. We represent this by defining specific conditions. Let us suppose that we define the condition $\omega = \{p1 \text{ wins the first set}\}$. Then, if a player must win two sets in order to win the game, it is easy to observe that if the $p1$ wins the first set, her probability for winning the match would be tend to increase regarding the probability at the start of the event. Thus, if $p1$ wins the first set, her new probability for winning the match is going to increase to:

$$(p|\omega = \text{true}) = \begin{cases} \rho' & \text{for } p1 \text{ to win the game} \\ (1-\rho') & \text{for } p1 \text{ not to win the game} \end{cases}$$

such that $\rho' > \rho$. Similarly, if the player does not win the first set, her probability for winning the match at the end of the first set is going to decrease to:

$$(p|\omega = \text{false}) = \begin{cases} \rho'' & \text{for } p1 \text{ to win the game} \\ (1-\rho'') & \text{for } p1 \text{ not to win the game} \end{cases}$$

such that $\rho'' < \rho$. Therefore, the value of odds in trading depends on the probabilities of increasing or decreasing while the trade is open and also on the probability of the ω condition being true.

3.2 Identifying Profitable Trades

If $p(\omega)$ is the probability for $p1$ to win the set and $p(\neg\omega)$ = 1-$p(\omega)$ is the probability for $p1$ not to win the set, the probability increase and decrease between the starting and the closing bet will occur in the long term as follows:

$$\begin{cases} (\rho'\text{-}\rho) & \text{Prob. of increasing with a probability of } p(\omega) \\ (\rho''\text{-}\rho) & \text{Prob. of decreasing with a probability of } p(\neg\omega) \end{cases}$$

Therefore, the decision to bet on *back* or on *lay* in the starting bet in order to make the opposite bet at closing is dependent on which of the following situations is fulfilled:

$$\begin{cases} (|\rho'\text{-}\rho| \times p(\omega)) > (|\rho\text{-}\rho''| \times p(\neg\omega)) & \text{Profit if starting is } \textit{back} \\ (|\rho'\text{-}\rho| \times p(\omega)) < (|\rho\text{-}\rho''| \times p(\neg\omega)) & \text{Profit if starting is } \textit{lay} \\ (|\rho'\text{-}\rho| \times p(\omega)) = (|\rho\text{-}\rho''| \times p(\neg\omega)) & \text{Fair odds} \end{cases}$$

It must be pointed out that in this example, if ω is true, this causes a probability increase. If the fulfillment of the ω condition makes a probability decrease (in this example could be that the $p1$ does not win the first set), the profit of the situations described above would be the inverse because in this case $\rho' < \rho$ and $\rho < \rho''$. Therefore, if we select a ω condition that makes a probability decrease when it is true, the trading decisions are defined as follows:

$$\begin{cases} (|\rho'\text{-}\rho| \times p(\omega)) > (|\rho\text{-}\rho''| \times p(\neg\omega)) & \text{Profit if starting is } \textit{lay} \\ (|\rho'\text{-}\rho| \times p(\omega)) < (|\rho\text{-}\rho''| \times p(\neg\omega)) & \text{Profit if starting is } \textit{back} \\ (|\rho'\text{-}\rho| \times p(\omega)) = (|\rho\text{-}\rho''| \times p(\neg\omega)) & \text{Fair odds} \end{cases}$$

3.3 Trading decisions

If a *starting* bet is placed at odds y regarding an outcome, and after δ time-steps a *closing* bet is placed at odds y^δ, the difference between both odds defines the *odds evolution* of this outcome for the next δ time-steps. This odds evolution is related as an opposite probability evolution for the next δ time-steps as:

$$\alpha(y)^\delta = \frac{1}{y^\delta} - \frac{1}{y}$$

The probability evolution represents a probability increase if the odds at the *closing* bet are lower than the odds at the *starting* bet; the probability evolution represents a probability decrease if the odds at the *closing* bet are higher than the odds at the *starting* bet:

$$\alpha(y)^\delta \begin{cases} < 0 & \text{if } y^\delta > y \\ > 0 & \text{if } y^\delta < y \end{cases}$$

Given a starting odd y, we can create a sequence of *probabilities evolution* $\alpha(y)^{\delta_1}, \alpha(y)^{\delta_2}, \ldots, \alpha(y)^{\delta_n}$, in which each element $\alpha(y)^{\delta_i}$ represents the predicted evolution after δ_i time-steps. Thus, this evolution represent how odds are expected to increase or decrease during the next few time-steps.

As we stated in Section 3.1, predictions for the next δ time-steps must be made depending on whether or not the ω condition is true or not:

$$\begin{cases} \alpha_t(y)^\delta & \text{if } \omega \text{ is true in the next } \delta \text{ time-steps} \\ \alpha_f(y)^\delta & \text{if } \omega \text{ is false in the next } \delta \text{ time-steps} \end{cases}$$

The probability of ω being true in the next δ time-steps allows us to define whether or not odds have value. Therefore, if the probability of ω being true in the next δ time-steps is $p(\omega^\delta)$ and the probability of being false is $p(\neg\omega^\delta)$ = (1-$p(\omega^\delta)$), the value of the odds y can be calculated according to the following cases.

The first case represents that the profits that could be obtained in the next δ time-steps are greater when ω is true than when ω is false:

$$\left(\left(p(\omega^\delta) \times \left|\alpha_t(y)^\delta\right|\right) - \left(p(\neg\omega^\delta) \times \left|\alpha_f(y)^\delta\right|\right)\right) > 0$$

If $\alpha_t(y)^\delta > 0$ means that odds are expected to decrease and therefore, the value is on betting on y at the *starting* bet, and betting against y at the *closing* bet after δ time-steps. If $\alpha_t(y)^\delta < 0$ means that odds are expected to increase and therefore, the value is on betting against y at the *starting* bet and betting on y at the *closing* bet.

The second case represents that the profits that could be obtained in the next δ time-steps are greater when ω is true than when ω is false:

$$\left(\left(p(\omega^\delta) \times \left|\alpha_t(y)^\delta\right|\right) - \left(p(\neg\omega^\delta) \times \left|\alpha_f(y)^\delta\right|\right)\right) < 0$$

If $\alpha_t(y)^\delta > 0$ means that odds are expected to increase and therefore, the value is on betting against y at the *starting* bet, and betting on y at the *closing* bet after δ time-steps. If $\alpha_t(y)^\delta < 0$ means that odds are expected to decrease and therefore, the value is on betting on y at the *starting* bet and betting against y at the *closing* bet.

Finally, if both terms are equal, we say that the odds are fair because there are no profits whatever the combination of bets is:

$$\left(\left(p(\omega^\delta) \times \left|\alpha_t(y)^\delta\right|\right) - \left(p(\neg\omega^\delta) \times \left|\alpha_f(y)^\delta\right|\right)\right) = 0$$

It must be pointed out that if y is related to a *back* bet, the inverse trading decision should appear for the *lay* bet.

Therefore, the major aim of taking trading decisions is to find the time-step δ_{max} that has the highest value for the odds:

$$\delta_{max} = argmax\left(\left|\left(p(\omega^\delta) \times \left|\alpha_t(y)^\delta\right|\right) - \left(p(\neg\omega^\delta) \times \left|\alpha_f(y)^\delta\right|\right)\right|\right)$$

This δ_{max} defines the time difference from the *starting* bet to the *closing* bet in order to maximize the profits.

4 The CBR reasoning mechanism

Sports betting markets represent a multilateral market model in which traders send their bets at their own odds to the mediator who matches compatible bets. Orders compete for the best *back* and *lay* offers. Therefore, the offers that cannot be matched remain waiting until they can be matched or are canceled. One of the tasks of the bookmaker is to also show at any time the best *back* and *lay* odds of these waiting offers. For a specific market, there is a list of all the *back* and *lay* bids that are currently waiting, which are ordered from the highest odds to the lowest.

In this work, we use Betfair[1] as the sports betting market studied. Betfair is the world's biggest prediction exchange. According to [8], Betfair processes more than 6 million transactions on an average day (more than all of the European stock exchanges combined). Betfair is based on the New York Stock Exchange model and allows players to bet at odds set by other players rather than the bookmaker.

The Continuous Double Auction (CDA) is a typical institution of real-world exchange markets, such as financial assets, foreign exchange, energy, etc. In this institution, buyers and sellers place their offers at any time. When a participant accepts a buy or a sell offer, a transaction is made. To model a sports betting market, we define a CDA institution where agents can interact to obtain information of the market at a given moment and where they can also place bets. In our model, Betfair acts as the mediator between users, matching the compatible bets and showing the best *back* and *lay* odds at a given time. As an interface of Betfair, we define the *bookmaker* agent, which acts as a gateway between Betfair and the agents. Thus, when an agent wants to request odds or wants to place offers, it needs to communicate with the *bookmaker* agent. If the agent sends offers, these will be matched by Betfair or will be queued in the waiting offers queue. If agents are requesting the current odds, the *bookmaker* agent will retrieve them by accessing Betfair. Therefore, from the point of view of other users, they do not know if they are trading with humans or agents.

The prediction model presented in [4], attempts to find repeated patterns of odd movements for different sporting events. This model captures some features of a current event and finds similarities with other past events by using an underlying CBR system. Then, we integrated this model into an agent, which by observing the odds evolution in these historical events, is able to predict the most accurate future odds depending on what happens during the event from that point on.

[1] http://www.betfair.com

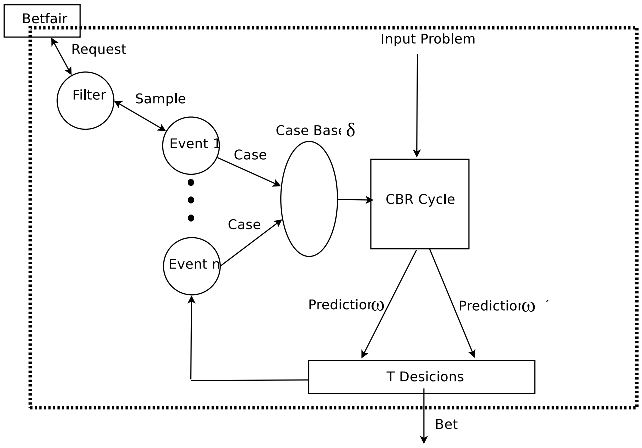

Fig. 1 CBR-Tagent

In this work, we extend the capabilities of this agent in order to provide it with trading decision capabilities by integrating the trading model described. Thus, the CBR-Trader agent (CBR-Tagent) attempts to simulate human behavior in sports betting markets. Humans base their trading decisions on their experience in the sports betting market. Through this experience, humans are able to know some odd movements and then, according to their expectations of odd movements, they can find value of odds. The CBR-Tagent bases its trading decisions on its odds prediction. Therefore, we can summarize the tasks of the CBR-Tagent (Figure 1) as follows:

- ***Data acquisition***: The first step is data acquisition according to the requirements of the problem. The CBR-Tagent interacts with the *bookmaker* agent in order to receive information about sport events.
- ***Creation of the case base***: From the data acquired, after a data filtering process (to exclude samples that may not reveal a real probability at a given moment), the CBR-Tagent creates the case base that will be used in the next task to predict future odds.
- ***CBR cycle***: Once the CBR-Tagent has created the case base, it is used to solve an unknown problem (in our case predict future odds) given similar past problems. This task represents the classic CBR cycle that was firstly introduced by Aamodt and Plaza in [2]. In our approach, the CBR-Tagent interacts with the *bookmaker* agent in order to obtain the information of a sporting event (unknown problem). Then, the CBR cycle measures the similarity between this problem and similar past problems. Finally, according to the future odds of the selected candidates, future odds are predicted for the unknown problem by adapting the solution.
- ***Trading decisions***: As the probability of an specific outcome is related to its odds, the CBR-Tagent predicts future odds regarding the outcome whether or not a specific condition is true. Then, according to the expected probability that this condition is true, the CBR-Tagent attempts to find value at different moments of the event, selecting the moment that takes the highest value for placing the *closing* bet.

In order to illustrate the tasks described in the previous section, we illustrate an example of trading in sports betting markets related to soccer events. We focus on trading on markets that are under/over 2.5 goals. These markets show the probability assessed by the participants for scoring less than 2.5 goals (0, 1, or 2) or more than 2.5 goals (3 or more) in a soccer match. First, the CBR-Tagent has to learn how the odds of the under/over markets evolve depending on the current features of a current game (data acquisition and creation of the case base tasks). Then, the CBR-Tagent predicts future odds by finding similarities with past events (CBR cycle). Finally, the CBR-Tagent uses its predictions to find the best trading decision based on the value for different moments of the events.

4.1 Data acquisition

We define the state of a soccer event at a specific moment according to the next properties:

- The exact moment of the game (in minutes).
- The current score of the game at that particular moment.
- The odds of the under/over markets. These show the *back* and *lay* odds for both the under and the over outcomes.
- The match odds. These show the *back* and *lay* odds for the home wins, visitor wins and draw wins.

The CBR-Tagent interacts with the *bookmaker* agent in order to obtain information from sporting events. This information is related to a specific event and it is stored as a sample. Each sample represents the state of a soccer event as a tuple $\langle m, s, h, v, d, u, o\rangle$, where:

- $0 \leq m \leq 45$ represents the minute of the game. For reasons of simplicity, we study the odds evolution in the first half of the event (45 minutes). Thus, this component is an integer. This restriction could be generalized to consider the second half of the events.
- $s \in \{0\text{-}0, 1\text{-}0, 0\text{-}1, 1\text{-}1, 2\text{-}0, 0\text{-}2\}$ represents the current score of the game. We take into account these current scores. Similarly than the minute component, this approach could be extended to consider other scores.

– h, v, and d, respectively, are the odds referring to home wins, visitor wins, and draw wins from the match odds market.
– u and o are the odds referring to the under outcome and the over outcome from the under/over markets.

Each h, v, d, u, and o has two real values $\langle b, l \rangle$ that represent the *back* and *lay* odds for the specific outcome.

Every 60 seconds, the CBR-Tagent requests these values from the *bookmaker*. Then, the CBR-Tagent creates a sequence of samples $x_1, x_2, x_3, \ldots x_n$ for a sporting event, ordered by the m component. To simplify notation, if $x_i = \langle m, s, h, v, d, u, o \rangle$, we write m_i to refer to m, and similarly for other components of x_i.

4.2 Creating the case base

In order to create the case base, the CBR-Tagent needs to represent information as a problem description and its solution. Each stored sample defines the state of the event at a specific moment. A pair of samples that represent the state of the event at two different moments, is used to create a case of the case base.

The problem description is each one of the stored samples of a single event, and the solution is the state of this event after δ minutes. Thus, given two different samples of the same event $x_i = \langle m_i, s_i, h_i, v_i, d_i, u_i, o_i \rangle$ and $x_j = \langle m_j, s_j, h_j, v_j, d_j, u_j, o_j \rangle$, such that $m_j - m_i = \delta$, we define a case of the case base as:

$$c^\delta = \langle m_i, s_i, h_i, v_i, d_i, u_i, o_i, u_j, o_j \rangle$$

This case represents the information of the event in the moment m_i (*starting* moment) and the information regarding the event after δ minutes, which in our example are the odds of the under/over outcomes: u_j and o_j.

Given different pairs of samples which their m components are different in δ minutes, we create a case base $C^\delta = \{c^\delta \mid c^\delta \text{ is defined}\}$. This case base stores cases of different events, but the information represented in a single case obviously refers to the same sporting event.

4.3 CBR cycle

The case base for a specific δ represents the information of events at a starting moment and their *back* and *lay* odds for the under/over markets in the next δ minutes. Therefore, given C^δ and an input problem $x = \langle m, s, h, v, d, u, o \rangle$, in which the *back* and *lay* odds of the under and over outcomes are $u = \langle b_u, l_u \rangle$ and $o = \langle b_o, l_o \rangle$, the CBR cycle predicts the *back* and *lay* odds for the under/over markets in the next δ minutes.

As each case is composed by several attributes, the matching process is divided in different steps in order to select the best candidates. First, an initial matching process retrieves a set of candidates based on the score components and the attributes that represent the odds that are predicted, in this example the *back* and *lay* odds for the under/over markets. Then, a more selective matching process is carried out to select the best candidates from the ones that were retrieved in the initial matching process, based on the minute component. This technique has been long ago mentioned by other authors [2]. For the first matching process, a similarity function is defined relative to the distance between the candidate and the input case. This similarity defines the distance between the four attributes that are considered (the *back* and *lay* odds for the under/over markets). The criteria used for retrieving cases in the initial matching process was to retrieve those candidates whose distance of each one of these four attributes are lower than a threshold τ with respect to the input case. As stated by Aamodt in [1], the threshold used in similarity functions can be gradually lower as more cases are added. In our case, a high threshold used in a large case base would cause that a lot of cases would be retrieved, while a low threshold in a small case base may cause that any case is retrieved.

For the selective matching process, a similarity function is defined relative to the time component between the candidate and the input case. Similar odds may evolve differently in the first minutes of the game than in the later ones. Therefore, the criteria used for retrieving cases in the selective matching process was to retrieve those candidates whose distance of the time component is lower than a threshold π with respect to the input case.

Finally, according to odds evolution of the cases that are retrieved, a solution for the input problem is constructed depending on the distance between each retrieved case and the input problem

The inference process of the CBR cycle can be summarized in the following steps:

Step 1. Retrieve the cases whose score components are the same as those of the input problem, and also whose *back* and *lay* odds for the under/over markets at the starting moment are the most similar to the odds of the input problem. For example, given an input problem $x = \langle m, s, h, v, d, \langle b_u, l_u \rangle, \langle b_o, l_o \rangle \rangle$ and a case of the case base $c_r = \langle m_r, s_r, h_r, v_r, d_r, \langle b'_u, l'_u \rangle, \langle b'_o, l'_o \rangle, u^\delta, o^\delta \rangle$, c_r is retrieved if the score is the same as the input problem and if its odds are not different from the odds of the input problem by more than the threshold τ, i.e. if the following formula is fulfilled:

$$s = s_r \wedge \mid b_u\text{-}b'_u \mid \leq \tau \wedge \mid l_u\text{-}l'_u \mid \leq \tau \wedge \mid b_o\text{-}b'_o \mid \leq \tau \wedge \mid l_o\text{-}l'_o \mid \leq \tau$$

As the accuracy of these odds is 0.02, the threshold τ was initially specified at 0.1 and has been adjusted up to 0.02 according to the case-base size increases.

Step 2. From all cases retrieved in Step 1, we select those whose time component is the most similar, given a threshold π. For example, given an input problem $x = \langle m, s, h, v, d, u, o \rangle$ and a case of the case base $c_r = \langle m_r, s_r, h_r, v_r, d_r, u_r, o_r, u^\delta, o^\delta \rangle$, we select those cases where $\mid m\text{-}m_r \mid < \pi$. As the accuracy of the time component is 1 minute, the threshold π was initially specified at 5 and has been adjusted according to the case-base size increases up to 1.

Step 3. From each case c_r selected in Step 2, where:

$$c_r = \langle m_r, s_r, h_r, v_r, d_r, \langle b_u, l_u \rangle, \langle b_o, l_o \rangle, \langle b'_u, l'_u \rangle, \langle b'_o, l'_o \rangle \rangle$$

for $r = \{1, 2 \ldots R\}$, being R the number of selected cases, we calculate the odds evolution for the under/over market as follows:

$$e(b,u) = (b_u - b'_u);\ e(l,u) = (l_u - l'_u)$$
$$e(b,o) = (b_o - b'_o);\ e(l,o) = (l_o - l'_o)$$

If the odds evolution is positive it means that the odds are going to decrease in the next δ minutes; if it is negative, the odds are going to increase in the next δ minutes. Then, we calculate an average odds evolution from all cases $c_1, c_2 \ldots c_R$, for the odds of the under/over market as follows:

$$A(b,u) = \frac{1}{R}\sum_{r=1}^{R} e(b,u)_r;\quad A(l,u) = \frac{1}{R}\sum_{r=1}^{R} e(l,u)_r$$

$$A(b,o) = \frac{1}{R}\sum_{r=1}^{R} e(b,o)_r;\quad A(l,o) = \frac{1}{R}\sum_{r=1}^{R} e(l,o)_r$$

Consider an input problem $x = \langle m, s, h, v, d, \langle b_u, l_u \rangle \langle b_o, l_o \rangle \rangle$ and the odds evolution for the under/over market: $A(b,u), A(l,u), A(b,o), A(l,o)$. Then, the predicted *back* and *lay* odds for the under/over markets are defined as: $predicted(p,k) = p_k + A(p,k)$ for each $p = \{b, l\}$ and $k = \{u, o\}$.

Step 4. If the predicted odds are finally similar to the real odds, the case is then retained in the case base. If one of the four predicted odds is different by more than a specified threshold than the real odds, the case is not retained, assuming that this case may be an anomalous case. If so, storing it could decrease the prediction accuracy of the entire system. This threshold has been specified similar to the τ threshold and we dynamically change this as the size of the case base increases in a similar way as we described above.

4.4 Trading decisions

Given an input problem $x = \langle m, s, h, v, d, u, o \rangle$ in which the *back* and *lay* odds for the under and over outcomes are:

$$u = \langle b_u, l_u \rangle \quad \text{and} \quad o = \langle b_o, l_o \rangle$$

the *back* and *lay* odds of the under and over outcomes in the next δ minutes are predicted as:

$$u^\delta = \langle b_u^\delta, l_u^\delta \rangle \quad \text{and} \quad o^\delta = \langle b_o^\delta, l_o^\delta \rangle$$

Therefore, the input problem represents the odds at the *starting* bet, and the predicted odds represent the odds expected at the *closing* bet. In order to the CBR-Tagent to be able to choose the best trading decision, it has to compute the probability increases and decreases depending on the odds evolution.

The difference between the odds y at the *starting* and *closing* bets, for each $y = \{b_u, l_u, b_o, l_o\}$, defines the *odds evolution* for the next δ minutes as we stated in Section 3.3.

$$\alpha(y)^\delta = \frac{1}{y^\delta} - \frac{1}{y}$$

for each component $y = \{b_u, l_u, b_o, l_o\}$.

Thus, given odds y at the *starting* moment, and the ω condition, we must calculate the minute δ_{max} that maximizes the value for the odds as we explained in Section 3.3

$$\delta_{max} = argmax \left(\left| \left(p(\omega^\delta) \times \left| \alpha_t(y)^\delta \right| \right) - \left(p(\neg\omega^\delta) \times \left| \alpha_f(y)^\delta \right| \right) \right| \right)$$

5 Evaluation

In order to evaluate the trading model proposed, in this section we present experiments which compare the performance of the CBR-Tagent and the performance of human traders. Eleven human volunteers with different levels of experience from the web page of sporting bets `http://foroapuestas.forobet.com/` were asked about different scenarios in real sports betting markets. We analyze the performance of the participants for predicting future odds and the performance for taking trading decisions. In Table 1 we show the level of experience in months for each human (Exp.) and their frequency of use of Betfair (Freq.), represented as weekly (W) and daily (D).

We use real data from soccer matches played during the 2009-2010 season in four of the most important leagues in the world: Premier League (England), Bundesliga (Germany), Primera Division (Spain), and Serie A (Italy). These competitions are selected because they are some of the most important leagues in the world, and therefore, sports betting markets regarding these leagues have a large number of traders. Thus, each

Table 1 Experience and frequency of use

	Exp.	Freq.
Human 1	>36	W
Human 2	3-12	D
Human 3	3-12	W
Human 4	>36	D
Human 5	>36	W
Human 6	>36	W

	Exp.	Freq.
Human 7	>36	W
Human 8	1-3	D
Human 9	3-12	W
Human 10	12-36	W
Human 11	12-36	W

event has a very high liquidity, which is important in order to obtain reliable results. For these experiments, we present the results related to the under 2.5 goals market.

We need to define the ω condition that has a probability associated of being true in the next δ minutes. For these experiments this condition is $\omega = \{$no goal is scored in the next δ minutes$\}$. This condition is selected because it causes a probability change, but other valid conditions could also be used. Depending on whether this condition is true or not, the CBR-Tagent predicts the increase or decrease of the under 2.5 goals probability in the next δ minutes according to the *starting* bet. Therefore, the CBR-Tagent calculates a probability for each value of δ. The values of delta used are $\delta = \{1, 5, 10, 15\}$, that is, the *closing* bet can be placed 1, 5, 10, or 15 minutes later from the *starting* bet.

Following we present the experiments evaluating two aspects: first we compare the odds prediction accuracy depending on whether or not ω is true, and second, we compare the performance for taking trading decisions.

The CBR-Tagent uses a case base size of 300 cases in order to make predictions. These cases are composed by 40 different matches, therefore, some of these cases are related to the same match at different moments. Although humans have experience in sports betting markets, we also provided them with information regarding these 40 matches that are used by the CBR-Tagent, in order to humans can see the odds evolution of these events depending on the ω condition. For testing the performance, we use 60 different matches that we divide in 3 sets of 20.

5.1 Odds prediction

The first experiment tries to evaluate the humans accuracy and the CBR-Tagent accuracy for predicting future odds depending on the ω condition.

First we evaluate the prediction of future odds for 20 matches at which no goal is scored in the next δ minutes, i.e., matches at which the ω condition is true. Thus, according to the state of the match at the *starting* bet, participants have to predict the future odds in the next 1, 5, 10 and 15 minutes. Table 2 shows the performance of humans and the CBR-Tagent in this experiment. We show the mean error rate (E) between the predicted odds and the real odds in cents, that is, how many cents the predicted odds are different from the real odds. We also show the standard deviation (σ) associated to this error rate.

Table 2 Odds prediction accuracy depending on when ω is true

	δ=1		δ=5		δ=10		δ=15	
	E	σ	E	σ	E	σ	E	σ
Human 1	2	1	2	2	5	5	9	8
Human 2	1	1	2	2	5	5	8	6
Human 3	3	2	5	3	5	4	7	5
Human 4	2	1	4	1	2	2	6	5
Human 5	4	1	6	4	3	2	4	3
Human 6	2	1	3	2	4	2	5	3
Human 7	3	2	5	3	7	6	8	7
Human 8	2	2	5	4	6	4	8	3
Human 9	2	1	4	2	5	4	7	5
Human 10	1	1	4	2	3	2	7	5
Human 11	3	1	5	2	7	6	10	7
CBR-Tagent	1	1	2	1	3	2	4	2

It must be pointed out that every participant correctly predicts odd movements in this experiment. Odds regarding the under outcome get lower if no goal is scored in the next δ minutes due to the probability of scoring less than 2.5 goals increases.

As can be observed, the error rate increases if δ increases. This means that it is more difficult to predict future odds when the time period is longer. However, these differences are not as high for the CBR-Tagent. It can be observed that the CBR-Tagent outperforms the behavior of every human in almost every situation. It is able to predict more accurate odds than humans with more than three years of experience in sports betting markets. Some humans have poor performance for predicting future odds, which may be due to their lack of experience. However, the most experienced humans do not make more accurate predictions than the CBR-Tagent. Humans can guide their predictions by their feelings, their emotions, the teams that are playing, or the players of the match. These factors are not taken into account by the CBR-Tagent, which shows a more rational behavior, and thus, better results.

The second experiment of odds prediction tries to evaluate the performance of the participants when the ω condition is false, that is, the performance for predicting future odds when a goal is scored. Humans and the CBR-Tagent predicts the future odds of 20 different events in which a goal is scored immediately. Participants have to predict the future odds of the under outcome by taking into account the current state of the match. Table 3 shows the performance of humans and the CBR-Tagent for predicting the odds when a goal is scored.

Table 3 Odds prediction accuracy depending on when ω is true

	E	σ
Human 1	13,5	9,5
Human 2	8,8	3,76
Human 3	18	7,1
Human 4	11,7	5,6
Human 5	9	5,5
Human 6	7,5	2,9

	E	σ
Human 7	11,3	3,1
Human 8	9,16	5,65
Human 9	10,1	5,14
Human 10	9,78	6,24
Human 11	14,36	8,45
CBR-Tagent	1,5	0,9

Similarly to the first experiment, each participant predicts correctly an odds increase, that is, the probability of scoring less than 2.5 goals decreases due to a goal is scored. However, this experiment shows greater differences between humans and the CBR-Tagent than the prediction when no goal is scored. As can be observed, the accuracy prediction of the CBR-Tagent for predicting future odds when a goal is scored is clearly higher than all the humans participating in the experiment, even the most experienced ones. In contrast to the first experiment, it is more difficult for humans to predict the specific future odds when a goal is scored than when a goal is not scored. This can be explained by the fact that when a goal is not scored, the odds tends to move slowly, and when a goal is scored, the odds increases suddenly causing the market to react abruptly to this change of condition. Humans assume that odds are going to increase, but they do not accurately predict how much this increase is. However, the accuracy of the CBR-Tagent is similar to the accuracy for predicting odds when no goal is scored. If the CBR-Tagent is able to make more accurate predictions than humans when the ω condition is true and when it is false, we can suppose that it can be easier for the CBR-Tagent to find the value of odds for trading than for humans. The next experiment attempts to evaluate this capability.

5.2 Trading decisions

We present an experiment related to trading decisions. In this experiment, humans and the CBR-Tagent have to take trading decisions according their expectations. We presented 20 scenarios related to specific states of matches, and we asked the participants about their trading decisions that they would take according to the state of the match. They must select whether to bet on *back*, bet on *lay*, or *not bet* (if the participant is reluctant to bet) at the *starting* bet (in the current moment). Then, according to the starting bet, they must select when to place the closing bet: after some minutes, when the odds reach some value, if a goal is scored, or *not bet*.

For the CBR-Tagent, the trading decision is dependent on the odds prediction and the ω condition. In order to assess the probability of the ω condition being true or not in the next few minutes, we take this probability as the percentage of matches from 2000 to 2010 in which ω was true in the next δ minutes according to the minute of the *starting* bet. As an example, if the number of matches where no goal was scored between the minute 20 and 21 was the 97%, this percentage was used as the probability of ω being true from the minute 20 and 21. Therefore, for each specific minute, there is an associated probability of the ω condition. Then, this probability is used to calculate the trading decision.

As we stated in Section 4, the trading decision taken by the CBR-Tagent takes into account the under probability increase or decrease according to whether or not ω is true depending on δ. This requires calculating the $\alpha_t(y)^\delta$ and the $\alpha_f(y)^\delta$ for each δ, as explained in Section 4.4. Then, according to the under probability increase or decrease in the next δ minutes, and the probability of ω being true or not in the next δ minutes, the CBR-Tagent chooses which δ is the most beneficial for placing the *closing* bet, and also the correct combination that maximizes the profits: whether to bet on *back* at the *starting* bet and to bet on *lay* in the *closing* bet, or the inverse.

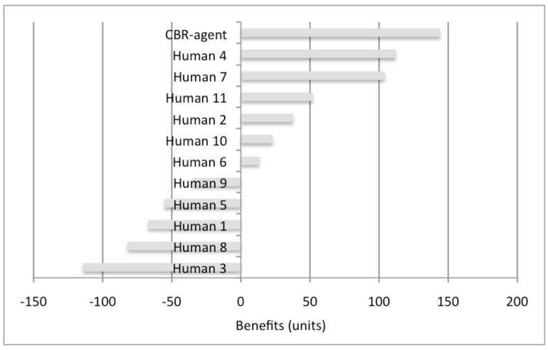

Fig. 2 Trading benefits

Figure 2 shows the performance of the humans and the CBR-Tagent after trading in these 20 scenarios. We present the results as a simulation related to the profits or losses in terms of units that each participant would make according to their trading decisions. We associate a linear stake of 100 units for each *starting* bet. As we can see, the CBR-Tagent is the participant which would win a large amount of money. As we might expect according to the odds prediction accuracy of the previous experiments, in general, the CBR-Tagent takes better trading decisions than humans, and this is translated into higher profits. Although it is not reflected in the figure, we can say that almost every human is reluctant to bet in about the 40-50% of the scenarios proposed. This means that they do not consider the scenario suitable to trade, i.e., they do not consider values in odds. However, the CBR-Tagent takes a trading decision in more than the 80% of the scenarios. This means that the capacity for finding value is is higher in the CBR-Tagent than in humans. Humans take into account factors such as the teams of the match, the team that is winning, the specific league of the match, etc.. This causes that humans choose not bet in more scenarios than the CBR-Tagent. Moreover, the CBR-Tagent does not take into account emotional factors and only considers whether or not a profit could be made in the long term according to the value of odds, without fear nor emotion. Therefore, the CBR-Tagent reasoning system can be considered to be more rational and powerful than humans behavior for this kind of trading scenarios.

As this percentage of matches where a trading is carried out influence the benefits obtained, we also show in Table 4, the yield obtained by participants. This yield is the internal rate of return and it is represented as the percentage of the benefits obtained according to the investments:

$$yield(\%) = \frac{benefits(units)}{investments(units)}$$

As we can observe, according to this parameter, the CBR-Tagent outperforms the behavior of every human. Nevertheless, the most performance humans (Human 4 and 7) have a yield very similar to the agent.

Table 4 Yield obtained by the participants

	Yield (%)
Human 1	-7
Human 2	4
Human 3	-10
Human 4	8.4
Human 5	-4.5
Human 6	1.1

	Yield (%)
Human 7	8.4
Human 8	-8
Human 9	-3.3
Human 10	1.6
Human 11	4.1
CBR-Tagent	8.5

As a conclusion for humans, the level of experience seems to influence the accuracy for prediction (Humans 5 and 6) as well as the final profit (Humans 4 and 7). Moreover, the worse results are caused by inexperienced traders such as the Human 3. This reveals that experience (approached as past cases), influence the reasoning system used by traders. It can also be observed that some humans are very prudent and, therefore, they usually take trading decisions in the short term, reducing the amount of profits that they could make if they were not so prudent. Other participants are riskier, and this human condition makes lose more than they might lose if they were not so risky. In contrast, the CBR-Tagent presents a behavior in which the average profit is similar to the average loss. This reveals that its decisions for the scenarios proposed are more rational than human decisions, because the CBR-Tagent is not as prudent as some humans, and is not as risky as others. This behavior has been more profitable in these experiments.

6 Discussion and Future Work

Most of the research on prediction markets is focused on pricing, that is, assessing the most accurate prices according to the probability of the event. Studies regarding prediction markets [12, 29] demonstrate that these markets provide very accurate probability predictions of future outcomes [15, 7, 30]. The pricing approach has been also studied for sports betting markets in [33]. However, as stated in [21], the shortcomings that plague individual decision makers could affect the behavior of a prediction market. Theoretical works also analyzes the complexity of scoring rules (such as Hanson or Logarithmic) for pricing [13]. Other works on prediction markets are focused on how information is incorporated into the market and therefore, how this information influences the prices. Pennock in [28] examines the political market reaction to influential news and how prices change. The work of Debnath et al. [10] is related to our work since they works on sports betting markets. They study how information regarding a sporting event affects prices. They show how the uncertainty regarding the correct outcome changes according to different states of the games. However, this work is different from our approach since they are not focused on the concept of trading.

Related to the probabilistic model presented in this paper, some other works can be found in the literature that provide other flexible variants to a general constrain-based CBR framework, such as probabilistic CBR frameworks [17]. In this paper, CBR techniques are approached together with probabilistic and fuzzy set-based modeling. Other approaches such as [14], also assess probabilities based on the frequencies of past cases. In [5], authors compare a probabilistic CBR approach with other well-known classification algorithms. This technique has been applied to other real-world domains such as collision avoidance systems [20].

To our knowledge, our work represents a novel work towards trading in prediction markets with the aim of making a profit regardless of the correct outcome. In order to carry out our work, we formulate the problem as predicting odds in sports betting markets. Other research areas such as financial or economic markets have studied the price evolution for predicting future prices. Since the introduction of computational tools for modeling financial and economic markets, most of the restrictions of classical analytic methods have been overcome. Several works related to stock markets have focused on modeling heterogeneous traders with different behaviors that can evolve over time [19,27]. Techniques from Artificial Intelligence such as rule-based systems or artificial neural networks have been broadly used for learning and evolving these artificial financial traders. Other works represent stock, commodity or foreign exchange markets as time series and study the evolution of the assets over time, trying to predict future prices [24, 16,32,18,3,22]. The price evolution in financial markets is dependent on external information whereas the odds evolution of sports betting markets is dependent on the outcomes probabilities. Thus, the evolution of both approaches can be studied from different techniques. We do not have proofs that the model presented in this paper for trading in sports betting markets would also have good performance for predicting price movements in financial markets. In addition, could also be interesting to apply techniques of financial markets to sports betting markets [31].

We approached the odds prediction on sports betting markets by studying the probabilities evolution. We based this technique on a simulation of the reasoning system similar to the one used by human traders. Human traders of sports betting markets take trading decisions according to their past experience in these markets. Based on this hypothesis, we have proposed a model for trading in sports betting markets based on a CBR technique. We have shown that experience influences the performance of human traders. Nevertheless, the CBR technique used in this paper offers a higher performance than humans. However, for future work we plan to compare the CBR technique used in this paper with other techniques used in other domains, such as the neural networks. According to the agent-based market design presented in this paper, we could develop agents based on other techniques and compare their performance.

In the Multiagent Systems area, several works have studied the performance of strategies in different trading scenarios. The Trading Agent Competition provides a scenario where traders exchange hotels, flights, or trips. Related works are pioneers in studying the problem of market odds prediction in a context of multi-auction environments [35]. Agent designers have come up with an interesting range of approaches for odds prediction in the context of the TAC market game. Our approach is similar to these works in the context of competitiveness. Sports betting markets are aimed to be competitive, and, therefore, traders with the highest performance will make the most profits. The Multiagent model presented for sports betting markets, provides us support for testing different trading strategies in these markets. Ontanon and Plaza [25,26] present an approach that combine Multiagent Systems and prediction markets. In this approach, the prediction market represents a social network in which agents can exchange arguments in order to predict new cases. One of the most interesting points of this approach is the use of an argumentation protocol inside the prediction market.

Experiments shown in this paper also provide us with information related to the behavior of human traders in sports betting markets. Emotional factors perceived by humans may influence their trading decisions, making them extremely prudent or risky. The non-emotional reasoning mechanism used by the CBR-Tagent has proved to be more profitable as we have shown in this work. The capability of the CBR-Tagent to find odds value is more accurate and rational than humans capability. One main reason for this behavior is that the CBR-Tagent is better than humans in predicting the changes in odds, as we have shown in the odds prediction when a goal is scored in the match. This make that the odds value can be found more accurate than in humans. Furthermore, real sports betting markets require to take decisions quickly, and therefore, human capability is also restricted by this factor. For future work, we also plan to apply this trading model to different markets in order to show whether or not the specific market influences the performance of the CBR-Tagent.

Acknowledgements This work has been partially supported by CONSOLIDER-INGENIO 2010 under grant CSD2007-00022, and project TIN2011-27652-C03-01. Juan M. Alberola has received a grant from Ministerio de Ciencia e Innovación de España (AP2007-00289).

References

1. A. Aamodt. Knowledge-intensive case-based reasoning and sustained learning. In *Topics in Case-based reasoning*, pages 274–288. Springer, 1990.
2. A. Aamodt and E. Plaza. Case-based reasoning; foundational issues, methodological variations, and system approaches. *AI COMMUNICATIONS*, 7(1):39–59, 1994.
3. J. J. Ahn, H. W. Byun, K. J. Oh, and T. Y. Kim. Bayesian forecaster using class-based optimization. *Applied Intelligence*, 36(3):553–563, 2012.
4. J. M. Alberola, A. Garcia-Fornes, and A. Espinosa. Price prediction in sports betting markets. In *Proceedings of the 8th German conference on Multiagent system technologies*, pages 197–208, 2010.

5. D. Arias-Aranda, J. L. Castro, M. Navarro, and J. M. Zurita. A cbr system for knowing the relationship between flexibility and operations strategy. In *Proceedings of the 18th International Symposium on Foundations of Intelligent Systems*, ISMIS '09, pages 463–472, 2009.
6. C. Ates. Prediction markets are only human: Subadditivity in probability judgments. In *MSC in Finance and International Business*, 2004.
7. M. Berlemann and C. Schmidt. Predictive accuracy of political stock markets - empirical evidence from a european perspective. Technical Report 2001-57, 2001.
8. Betfair. http://www.betfaircorporate.com, 2009.
9. Y. Chen, S. Goel, and D. Pennock. Pricing combinatorial markets for tournaments. In *STOC '08: Proceedings of the 40th annual ACM symposium on Theory of computing*, pages 305–314. ACM, 2008.
10. S. Debnath, D. M. Pennock, C. L. Giles, and S. Lawrence. Information incorporation in online in-game sports betting markets, 2003.
11. B. Fischoff, P. Slovic, and S. Lichtenstein. Knowing with certainty: The appropriateness of extreme confidence. In *Journal of Experimental Psychology: Human Perception and Performance*, pages 552–564, 1997.
12. R. Forsythe, T. Rietz, and T. Ross. Wishes, expectations and actions: a survey on price formation in election stock markets. *Journal of Economic Behavior & Organization*, 39(1):83–110, 1999.
13. L. Fortnow, J. Kilian, D. Pennock, and M. Wellman. Betting boolean-style: A framework for trading in securities based on logical formulas, 2003.
14. G. Gayer. Perception of probabilities in situations of risk: A case based approach. *Games and Economic Behavior*, 68(1):130–143, 2010.
15. M. Guo and D. Pennock. Combinatorial prediction markets for event hierarchies. In *Proc. of The 8th AAMAS '09*, pages 201–208. Int. Foundation for Autonomous Agents and Multiagent Systems, 2009.
16. W. Huang, K. Lai, Y. Nakamori, and S. Wang. Forecasting foreign exchange rates with artificial neural networks: A review. *International Journal of Information Technology and Decision Making*, 3(1):145–165, 2004.
17. E. Hüllermeier. *Case-Based Approximate Reasoning*, volume 44 of *Theory and Decision Library*. Springer, 2007.
18. K.-J. Kim and H. Ahn. Simultaneous optimization of artificial neural networks for financial forecasting. *Applied Intelligence*, 36(4):887–898, 2012.
19. B. LeBaron. Agent based computational finance: Suggested readings and early research. Journal of Economic Dynamics and Control, 1998.
20. Y. Liu, C. Yang, Y. Yang, F. Lin, X. Du, and T. Ito. Case learning for cbr-based collision avoidance systems. *Applied Intelligence*, 36(2):308–319, 2012.
21. B. C. Love. Behavioural finance and sports betting markets. In *MSC in Finance and International Business*, 2008.
22. C. Luque, J. M. Valls, and P. Isasi. Time series prediction evolving voronoi regions. *Applied Intelligence*, 34(1):116–126, 2011.
23. R. L. D. Mantaras, D. McSherry, D. Bridge, D. Leake, B. Smyth, S. Craw, B. Faltings, M. Maher, C. Lou, M. C. K. Forbus, M. Keane, A. Aamodt, and I. Watson. Retrieval, reuse, revision and retention in case-based reasoning. *Knowl. Eng. Rev.*, 20(3):215–240, 2005.
24. J. Moody. Economic forecasting: Challenges and neural network solutions. In *In Proceedings of the International Symposium on Artificial Neural Networks*, 1995.
25. S. Ontañón and E. Plaza. Argumentation-Based Information Exchange in Prediction Markets. *Argumentation in Multi-Agent Systems*, 5384:181–196, 2009.
26. S. Ontañón and E. Plaza. An argumentation framework for learning, information exchange, and joint-deliberation in multi-agent systems. *Multiagent and Grid Systems Journal*, 7:95–108, 2011.
27. R. Palmer, W. Arthur, J. Holland, B. Lebaron, and P. Tayler. Artificial economic life: a simple model of a stock market. *Physica D*, 75:264–274, 1994.
28. D. Pennock, S. Debnath, E. Glover, and C. Giles. Modelling information incorporation in markets, with application to detecting and explaining events. In *Proceedings of the 18th Annual Conference on Uncertainty in Artificial Intelligence (UAI-02)*, pages 405–4, San Francisco, CA, 2002. Morgan Kaufmann.
29. D. Pennock, S. Lawrence, F. Nielsen, and C. Giles. Extracting collective probabilistic forecasts from web games, 2001.
30. C. Plott. Markets as information gathering tools. pages 1–15, 2000.
31. B. Qian and K. Rasheed. Stock market prediction with multiple classifiers. *Applied Intelligence*, 26(1):25–33, 2007.
32. S. Raudys and I. Zliobaite. The multi-agent system for prediction of financial time series. In *ICAISC*, volume 4029, pages 653–662. Springer, 2006.
33. C. Schmidt and A. Werwatz. How accurate do markets predict the outcome of an event? the euro 2000 soccer championship experiment, 2002.
34. S. C. K. Shiu and S. K. Pal. Case-based reasoning: Concepts, features and soft computing. *Applied Intelligence*, 21(3):233–238, 2004.
35. M. P. Wellman, D. M. Reeves, K. M. Lochner, and Y. Vorobeychik. Price prediction in a trading agent competition. *Journal of Artificial Intelligence Research*, 21:2004, 2004.