


Article

Assessing the Relevance of Opinions in Uncertainty and Info-Incompleteness Conditions

Gerardo Iovane ^{1,†}, Riccardo Emanuele Landi ^{2,*}, Antonio Rapuano ^{1,†}  and Riccardo Amatore ^{1,†}

¹ Department of Computer Science, University of Salerno, 84084 Fisciano, Italy; giovane@unisa.it (G.I.); arapuano@unisa.it (A.R.); ramatore@unisa.it (R.A.)

² Rigenera S.r.l., Via Aventina 7, 00153 Rome, Italy

* Correspondence: riccardo.landi@rigenera2020.it

† These authors contributed equally to this work.

Abstract: Researchers are interested in defining decision support systems that can act in contexts characterized by uncertainty and info-incompleteness. The present study proposes a learning model for assessing the relevance of probability, plausibility, credibility, and possibility opinions in the conditions above. The solution consists of an Artificial Neural Network acquiring input features related to the considered set of opinions and other relevant attributes. The model provides the weights for minimizing the error between the expected outcome and the ground truth concerning a given phenomenon of interest. A custom loss function was defined to minimize the Mean Best Price Error (MBPE), while the evaluation of football players' was chosen as a case study for testing the model. A custom dataset was constructed by scraping the Transfermarkt, Football Manager, and FIFA21 information sources and by computing a sentiment score through BERT, obtaining a total of 398 occurrences, of which 85% were employed for training the proposed model. The results show that the probability opinion represents the best choice in conditions of info-completeness, predicting the best price with 0.86 MBPE (0.61% of normalized error), while an arbitrary set composed of plausibility, credibility, and possibility opinions was considered for deciding successfully in info-incompleteness, achieving a confidence score of 2.47 ± 0.188 MBPE ($1.89 \pm 0.15\%$ of normalized error). The proposed solution provided high performance in predicting the transfer cost of a football player in conditions of both info-completeness and info-incompleteness, revealing the significance of extending the feature space to opinions concerning the quantity to predict. Furthermore, the assumptions of the theoretical background were confirmed, as well as the observations found in the state of the art regarding football player evaluation.

Keywords: decision support systems; uncertainty; info-incompleteness; machine learning; artificial intelligence; football market; athlete evaluation



Citation: Iovane, G.; Landi, R.E.; Rapuano, A.; Amatore R. Assessing the Relevance of Opinions in Uncertainty and Info-Incompleteness Conditions. *Appl. Sci.* **2022**, *12*, 194. <https://doi.org/10.3390/app12010194>

Academic Editors: João M. F. Rodrigues, Pedro J. S. Cardoso and Marta Chinnici

Received: 25 November 2021

Accepted: 23 December 2021

Published: 25 December 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 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 (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The main concept used for estimating the possibility of the occurrence of an event is the probability, in which the certain event is the upper extreme and the impossible event is the lower one. Between these two bounds, there are more or less probable events. Probability, however, is a limited concept, as it can be affected significantly by the lack of information; for instance, in the case that all the information on the environment in which dice are thrown is known, it is possible to easily decide the exact face of the dice that will be obtained. However, there are some events in which the probability can be computed through some additional information; i.e., in a dice throw, the number of faces can be known, while, in a financial context, where it is intended to compute the probability of an asset reaching a certain price, the information is almost totally absent. This is called an *uncertain and info-incomplete* environment.

To better estimate the occurrence of an event, different definitions of what is called *plausible reasoning* have been proposed. Iovane et al. [1] have modeled decision making and

reasoning in uncertainty and info-incompleteness conditions as the evaluation of *Probability*, *Plausibility*, *Credibility*, and *Possibility*, providing several models of interest (capital letters are used for specifying the concepts as defined by the authors). Probability is conceived as an estimate of evidence concerning a given phenomenon and Plausibility, Credibility, and Possibility as opinions extracted from the area of knowledge which does not regard direct evidence. In fact, in the case the probability estimation (i.e., the direct evidence) that a phenomenon occurs is weak, it is possible to reduce the uncertainty by acquiring information concerning, e.g., the opinion of experts or the sentiment of a group of people regarding the aforementioned phenomenon. When an estimate of evidence is available, i.e., a probability, we decide in conditions of *info-completeness*; when an estimate of evidence is not available, we decide in conditions of *info-incompleteness*. In a nutshell, when an event is very unlikely, or a decision based on the analysis of evidence cannot be performed, it is more promising to consider other sources of information, which may be less reliable than deciding “blindly”.

In the present work, it is proposed to re-enforce the model defined by Iovane et al. [1] by using machine learning to estimate the relevance of Probability, Plausibility, Credibility, and Possibility opinions in conditions of both info-completeness and info-incompleteness. To achieve the above goal, it was decided to adopt, as we explain in the following Sections, the *best price* model, as proposed by the authors, to a real study case. In particular, we will refer to the football players’ market, where each athlete is characterized by an economic evaluation. A custom dataset was built to train and test an Artificial Neural Network (ANN) in estimating the weights of Probability, Plausibility, Credibility, and Possibility in the above field of interest.

The work is organized as follows. In Section 2, a summary of the prodromic theory [1] is provided. Section 3 analyzes the state of the art and the most important studies in the field of decision making and reasoning in conditions of uncertainty and info-incompleteness, as well as in the research regarding the evaluation of athletes through Artificial Intelligence methodologies. Section 4 shows our proposed solution for weighting the opinions, while Section 5 describes the dataset and the implementation of the proposed learning model on the chosen study case. Finally, Section 6 discusses the results, and Section 7 summarizes the work and indicates the future direction.

2. Theoretical Background

In this section, a discussion of the advancements in the plausibility theory, with particular regard to the prodromic study [1], is provided.

According to Polya [2,3], the plausible reasoning is not subjective, and it is treated as a conditional probability: given two events A and B, the author conceives the plausibility as the confidence of B given that A is true. After Polya, Dempster–Shafer’s theory was defined [4,5]: the plausibility is no longer intended as one-dimensional but as a series of mutually exclusive alternatives with a maximum probabilistic value. To overcome the difficulties of Dempster–Shafer’s model, a new solution, called Dezert–Smarandache’s theory, was proposed [6]: there, the *plausibility* becomes an upper limit of the probabilistic value concerning a given event. Iovane et al. [1] provided a further contribution by extending the concept of plausibility to *credibility* and *possibility*. The authors defined the expectation function obtained through the composition of Probability, Plausibility, Credibility, and Possibility.

The study provides seven models for computing the above function. By considering $P_1, P_2, P_3, P_4 \in \mathbb{R}^k$, with $k > 0$, as the opinions concerning Probability, Plausibility, Credibility, and Possibility, respectively, the aforementioned models can be summarized as follows.

- *Average model*: the simplest model. It computes the average of the four opinions as

$$a_1 = a_1(P_1, P_2, P_3, P_4) = \frac{1}{4} \sum_{i=1}^4 P_i. \quad (1)$$

The model assigns the same importance to the different distributions of Probability, Plausibility, Credibility, and Possibility.

- *Product model*: the expectation function is defined by the P_i product as

$$a_2 = a_1(P_1, P_2, P_3, P_4) = \prod_{i=1}^4 P_i. \tag{2}$$

- *Weighted average model*: a_1 and a_2 assume that all P_i have the same importance. Instead, this model extends a_1 by weighting the P_i s with

$$\sum_{i=1}^4 \alpha_i = 1, \tag{3}$$

where α_i is the weight of P_i and a_3 is defined as

$$a_3 = \sum_{i=1}^4 \alpha_i P_i. \tag{4}$$

- *Weighted product model*: extends a_2 weighting the P_i s as well. Formally,

$$a_4 = \prod_{i=1}^4 \alpha_i P_i. \tag{5}$$

- *Overlap model with shift based on probability*: differently from other previously defined models, the overlap with shift based on probability allows a hierarchical use of the P_i s. Formally,

$$a_5 = \begin{cases} P_1 & \text{if } 1\% < P_r \leq 100\%, \\ P_2 & \text{if } 0.1\% \leq P_r \leq 1\%, \\ P_3 & \text{if } 0.01\% \leq P_r \leq 0.1\%, \\ P_4 & \text{if } P_r \leq 0.01\%. \end{cases} \tag{6}$$

In this model, the expectation function is selected from Probability, Plausibility, Credibility, and Possibility. There, the selection depends on the classical probability value (P_r). Each P_i , if selected, has a coefficient

$$0\% \leq c_i \leq 100\%. \tag{7}$$

- *Overlap model with shift based on hierarchical P_i* : the probability does not have a pivotal role; this is an alternative model for a_5 . There, the expectation function is defined by the authors as

$$a_6 = \begin{cases} P_1 & \text{if } \bar{P}_{P_1} - 3\sigma_{P_1} \leq P_r \leq \bar{P}_{P_1} - 3\sigma_{P_1}, \\ P_2 & \text{if } \bar{P}_{P_2} - 3\sigma_{P_2} \leq P_r \leq \bar{P}_{P_2} - 3\sigma_{P_2} \text{ and } P_r > \bar{P}_{P_2} - 3\sigma_{P_1}, \\ P_3 & \text{if } \bar{P}_{P_3} - 3\sigma_{P_3} \leq P_r \leq \bar{P}_{P_3} - 3\sigma_{P_3} \text{ and } P_r > \bar{P}_{P_2} - 3\sigma_{P_2}, \\ P_4 & \text{if } \bar{P}_{P_4} - 3\sigma_{P_4} \leq P_r \leq \bar{P}_{P_4} - 3\sigma_{P_4} \text{ and } P_r > \bar{P}_{P_3} - 3\sigma_{P_3}. \end{cases}$$

There, i.e., in a financial context like in the present work, in the case the goal is to obtain the best price of an athlete, \bar{P} represents the average price in the dataset and σ the standard deviation. It is important to note that, in this context, the above P_i s identify the distribution of \bar{P} and σ and not the P_i value. In other words, the selection of the P_i depends on the values of σ and \bar{P}_i and the P_i distribution.

- *Model based on Dempster's composition rules:* the last model the authors proposed is based on the Dempster's composition rules, which are defined only for the plausibility. Iovane et al. [1] extended those rules to Probability, Plausibility, Credibility and Possibility. Formally,
 - $m(P_1)$, relative $bel(P_1)$ and $Dpl(P_1)$;
 - $m(P_2)$, relative $bel(P_2)$ and $Dpl(P_2)$;
 - $m(P_3)$, relative $bel(P_3)$ and $Dpl(P_3)$;
 - $m(P_4)$, relative $bel(P_4)$ and $Dpl(P_4)$;

where m is called "mass function" by Dempster (the degree of belief), bel is the belief function, and Dpl represents the plausibility.

While the described models are used to compute the most likely price, i.e., the best price, we can extend the information defining the *occurrence* as

$$O(E) = \alpha_1 P_1(E) + \alpha_2 P_2(E) + \alpha_3 P_3(E) + \alpha_4 P_4(E), \quad (8)$$

where $P_i(E) : \mathbb{R}^k \rightarrow [0, 1]$, with $k > 0$, is how probable the event E is and α_i represents the weights of the $P_i(E)$, with $0\% \leq \alpha_i \leq 100\%$. Once the best price and how much this price occurs are determined, the last thing needed for describing an event is the reliability of the information. To compute the reliability of the best price, we can use the standard deviation. Formally,

$$R(E) = \sigma^2(E) = \sqrt{\frac{\sum_{i=1}^N (x_i - \bar{x})^2}{N - 1}}, \quad (9)$$

where x_i is in the set P_1, P_2, P_3, P_4 .

Therefore, the final output of the model is the triad

$$(E, O(E), R(O)). \quad (10)$$

In the prodromal study [1], the authors simulated the datasets to prove the correctness of the models. In the present work, it is intended to face a real case by using a neural network for estimating the weights of the best price associated with the expectation function a_3 . In the next section, the state of the art regarding athletes' price estimation is analyzed.

3. Related Work

The economical evaluation of football players is a much-addressed issue. In particular, in the financial area, the evaluation of an asset is made by supply and demand. The financial world applied to the sport is complex; differently from traditional finance, there are only two actors in the negotiation of a player: the buyer and the seller. They can agree on any price, and this can lead to several problems from a regulatory point of view. The question in this field is: can we have a reference point for the football players' evaluation? In this context, crowdsourcing is significant through Transfermarkt, but a more reliable tool is still needed.

As the financial world behind football, as well as the sports world in general, is vast and based on complex economic models, several studies have investigated how to predict or estimate the athletes' market value. Dobson and Goddard [7] proposed an interesting and detailed study concerning the economics of professional English football at the club level. As mentioned in the previous section, Iovane et al. [1] applied the described models to two different applications to prove the validity of the theory. The two experiments consisted of two simulations regarding the probabilities fields of biometrics and sport odds; there, the authors simulated the datasets, adding uncertainty through randomness over the input space. The weights α_i were defined without a backtest; thus, a deep study on the estimation of the weights is needed. The present work is conceived to solve the above difficulty.

In [8,9], the authors tried to estimate the market value of football players. Behravan and Razavi [8] clustered the football players by roles; after that, they used a hybrid regression method involving Particle Swarm Optimization (PSO) and Support Vector Regressor (SVR) for each cluster. They obtained a final accuracy for their model of 74%.

Furthermore, the sentiment can affect the athlete evaluation: an interesting study was conducted by Singh and Lamba [10]; they described how crowdsourcing, previous year statistics, and popularity of players can affect the evaluation. Regarding crowdsourcing, in [11], the authors proved how, in the context of German soccer, a community became the main source for reporting market values to predict the actual transfer fees. The authors described the evaluation process performed by the community, together with the accuracy of the estimated market values, and which variables are important to make a price estimation. They found that the variables that are mostly correlated with the price are those of age, precision, success, assertion, and flexibility.

The importance of athletes' age is analyzed and discussed in several papers. In particular, Gonzalez et al. [12] investigated the relative age effect, which was predominant in players born in the first months of the year compared to those born in the last months. The results show that, except for the youth categories, the relative age does not affect the professional football player market evaluation but only the selection in youth categories.

Other scholars investigated the variables affecting the football players' market value. In [13], the values of Transfermarkt.de were acquired, while in [14], the authors analyzed the Football Manager game values. Felipe et al. [15] investigated the influence of team variables and the player role on the athletes' market values. Another interesting variable for the price estimation of football players is popularity; Franck and Nüesch [16], as well as Kiefer [17], investigated the influence of the players' popularity on their market value. The authors proved that the market value of the players is influenced by both talent and non-performance-related popularity. In [18], the authors proposed a decision support system for football club managers and players' agents by estimating the correct wages of football players. Player skills, performances in the previous season, age, the trajectory of the improvement, personality, and other features were considered.

Although there are several works that estimate the economical value of athletes, a decision support system merging different opinions can perform well even in conditions of uncertainty and info-incompleteness. The present work aims to investigate the roles and the weights of Probability, Plausibility, Credibility, and Possibility in the market evaluation of football players, for both improving the state of the art in the research area and providing a case study in which the assessment of the best opinion, in conditions of info-incompleteness, can be achieved.

4. Proposed Solution

Referring to the best price function defined in (3), it was intended to predict the weights α_1 , α_2 , α_3 , and α_4 , related to the opinions associated with Probability, Plausibility, Credibility, and Possibility, respectively, by extracting data from different sources. Given the occurrences of the above four opinions and the related ground truth, it is possible to approximate the best price function, i.e., the best opinion, by defining a learning model trained to predict the weights. This approach permits obtaining the relevance of opinions in conditions of both info-completeness and info-incompleteness, together with the most promising estimate regarding a given phenomenon.

To achieve the above goal, it was decided to propose the model described in Figure 1, in which an ANN receives four features associated with Probability, Plausibility, Credibility, and Possibility opinions and an arbitrary set of other relevant attributes. Each of the opinions is extracted from a dedicated source of information, as well as the additional set of attributes.

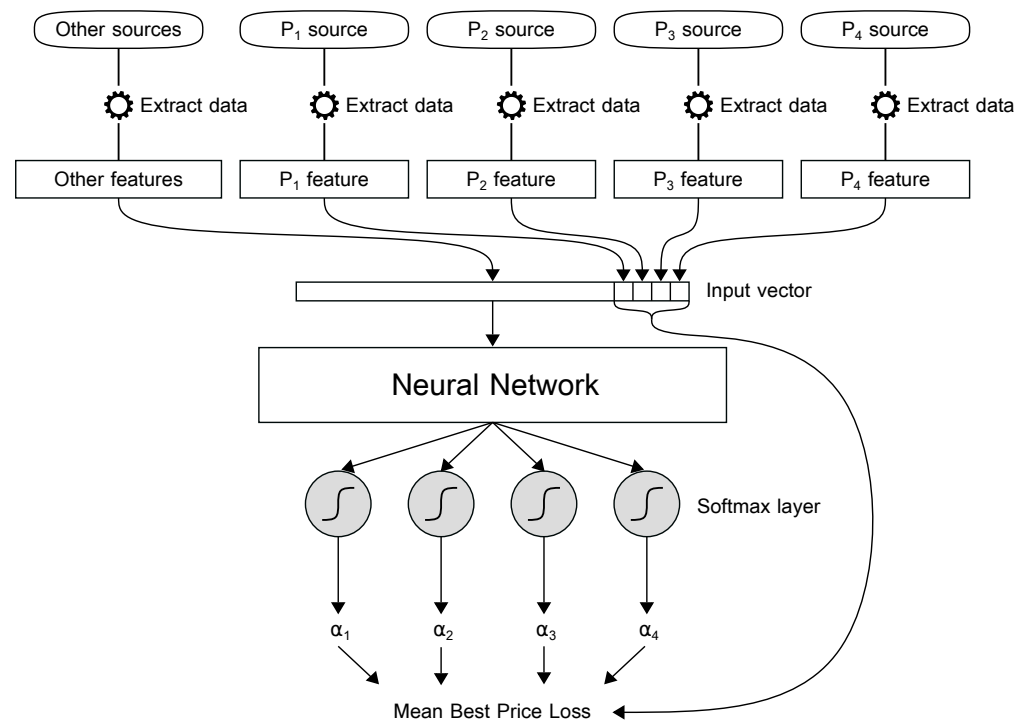


Figure 1. The proposed learning model to predict the relevance of opinions and an estimate of the best opinion in uncertainty and info-incompleteness conditions.

The model outputs four weights extracted from the components of a softmax layer to minimize a custom loss function, called *Mean Best Price Loss*, which was defined as

$$L_{MBP} = \frac{1}{N} \sum_{i=1}^N |P_{0_i} - (\alpha_{1_i}P_{1_i} + \alpha_{2_i}P_{2_i} + \alpha_{3_i}P_{3_i} + \alpha_{4_i}P_{4_i})|, \tag{11}$$

where N is the batch size, while P_{0_i} , P_{1_i} , P_{2_i} , P_{3_i} , and P_{4_i} are the ground truth and the Probability, Plausibility, Credibility, and Possibility opinions concerning the i -th occurrence, respectively. Similarly, the weights α_{1_i} , α_{2_i} , α_{3_i} , and α_{4_i} are the components of the softmax layer concerning the i -th occurrence. The loss function can be considered as the Mean Absolute Loss, in which the second term of the subtraction represents the prediction of the proposed model given a certain instance of features. The performance metric associated with the above loss is called *Mean Best Price Error* (MBPE).

The proposed approach above permits the ANN to learn the parameters that provide the optimal weights for minimizing the difference between the best price prediction, i.e., the estimate of the best opinion, and the true expected outcome. For instance, suppose there is some interest in evaluating the reliability in the occurrence of a given phenomenon. Suppose the phenomenon is also characterized by some specific attributes and that there exist sources of information from which one or more opinions can be extracted. In the case evidence about the phenomenon exists, a probable occurrence can be obtained; in the case some experts are involved in the study of the phenomenon, a plausible occurrence can be obtained; in the case the people discuss the phenomenon, a credible occurrence can be considered; in the case some other less relevant sources of information are available, a possible occurrence can be extracted. The acquisition of opinions related to a common domain of attributes permits an inference to be performed on the occurrence related to the given phenomenon of interest.

Under the above hypotheses and definitions, we have decided to test the proposed approach on a real study case regarding football players’ market evaluation. Such a problem is ideal for studying the relevance of opinions in uncertainty and info-incompleteness

conditions, as the prediction of the next transfer cost of a player is subjected to several sources of speculation. It is possible to evaluate the proposed model in conditions of both info-completeness, i.e., when the opinion concerning the Probability is available, and info-incompleteness, i.e., when the opinion related to the Probability is completely lacking. The following section evaluates the MBPE error as the subsets of opinions vary; for evaluating the performance of the model in conditions of info-incompleteness, e.g., the feature related to the Probability is set to zero, so as to test the case in which a decision should be taken when only Plausibility, Credibility, and Possibility opinions are fully or partially available.

5. Experiments on a Study Case: Football Players Evaluation

As discussed in the previous section, to provide a case study on which to verify the effectiveness of the proposed model, we performed the prediction of the transfer cost of a player starting from their attributes and the opinions associated with the Probability, Plausibility, Credibility, and Possibility, extracted from the web. Formally, in the present use case the opinions are defined as $P_1, P_2, P_3, P_4 \in \mathbb{R}$, since prices are one-dimensional quantities.

To realize the above scope, four sources of information were chosen:

1. Transfermarkt players evaluation in the year 2021/2022 for the opinion related to Probability;
2. Football Manager players evaluation for the opinion related to Plausibility;
3. Wikipedia descriptions as a sentiment, combined with Football Manager society costs of players, for the opinion related to Credibility;
4. FIFA21 players evaluation for the opinion related to Possibility.

Each source of information is normalized as a price, e.g., having a player with Transfermarkt, Football Manager, and FIFA21 evaluations of 16, 15, and 25 million of euros for the Probability, Plausibility, and Possibility opinions, respectively, while a Football Manager society has a cost of 20 million euros, combined with a sentiment score of 0.95, for the Credibility opinion. Data extraction was performed by scraping, through four distinct scripts, the web pages related to the major European football leagues, i.e., Serie A, La Liga, Premier League, Ligue 1, and Bundesliga. The ground truth was found by acquiring the cost of transfers from the same source chosen for finding the opinions related to Probability. Regarding the players' attributes, it was decided to acquire their characteristics from the same source adopted for finding the opinions related to Possibility.

The data extraction process can be summarized, for each player, as the parallel execution of the following tasks:

- Extraction of the Transfermarkt evaluation, at one year before the next transfer, to obtain the Probability feature P_{1i} ;
- Extraction of the Football Manager evaluation, at one year before the next transfer, to obtain the Plausibility feature P_{2i} ;
- To obtain the Credibility feature P_{3i} , performing the computation of a sentiment score from the related Wikipedia description, at one year before the next transfer, and weighting the result to the extracted Football Manager society cost;
- Extraction of the FIFA21 evaluation, at one year before the next transfer, to obtain the Possibility feature P_{4i} ;
- Extraction of the FIFA21 attributes, at one year before the next transfer, to obtain other features.

In the following sections, a custom dataset composed of data extracted from the different sources of information and the adopted feature selection methodology, together with the discussion of the experimental results, is presented. It was decided to investigate the proposed model in conditions of info-completeness, i.e., when the opinion related to Probability is considered, and under different configurations of uncertainty and info-incompleteness, i.e., when the opinions related to Plausibility, Credibility, and Possibility are completely or partially available.

5.1. Sentiment Analysis

The analysis of sentiment was conducted on the texts acquired from the Wikipedia pages concerning the players. The text sentiment classification model employed is based on BERT (Bidirectional Encoder Representations from Transformers) [19] and fine-tuned on the IBM Claim Stance Dataset [20,21]. The solution receives a text string as input, while it outputs a score in the range $[0, 1]$, in which 0 and 1 represent the most negative and positive emotions, respectively. The BERT framework permits fine-tuning a pre-trained language model for tackling several Natural Language Processing tasks, of which in the present study the sentiment analysis was valued. The model is characterized by two operational phases: (i) *pre-training*, for training, through an unsupervised approach, the model over several tasks; (ii) *fine-tuning*, for optimizing, through a supervised approach, the parameters found in the previous task to the sentiment analysis. For both processes, the same Multi-layer Bidirectional Transformer Encoder architecture was employed: it generates word embeddings, i.e., univocal probabilistic representations of words, through a bi-directional training approach; for the first task, i.e., pre-training, the network is trained for solving the two problems of Masked Language Modeling (MLM) and Next Sentence Prediction (NSP); for the second task, i.e., fine-tuning, a custom fully connected layer is added to the pre-trained network for solving the desired supervised problem. The MLM problem consists in predicting the masked words of a sentence to learn their bi-directional context, while the NSP problem regards the determination of the order through which two sentences are employed in a given context.

The classifier reaches, in terms of accuracy, 94% and provides a good estimate of the sentiment in a text. In the present study, the output of the aforementioned model is used as an opinion modulation factor characterizing the sentiment associated with a given phenomenon.

5.2. Dataset

The data were extracted from the Transfermarkt, FIFA21, and Football Manager sources available on the web. In particular, Transfermarkt was the first source explored, as the extraction of the data was made dependent on the latest transfers of the calendar year 2021. The employed scraping process extracts the names and the related market values and transfer costs from Transfermarkt; then, for each extracted name, the sources related to FIFA21 and Football Manager are considered for obtaining attributes and values concerning Plausibility, Credibility, and Possibility opinions. Meanwhile, Wikipedia pages regarding the extracted names are processed through the sentiment analysis algorithm described in Section 5.1. The players' transfer costs provided by Transfermarkt (min = 0.1 million of euros, max = 125 million of euros) serve as ground truth for the training and test processes. A total of 398 data occurrences of transferred players have been considered.

Figure 2 shows the pair distributions of Probability, Plausibility, Credibility, and Possibility features obtained by building the dataset.

The pair distributions enhance the characteristics in the population of opinions by considering two features at a time as coordinates. The Figure, on the first diagonal, also shows the density distributions related to the opinions. It can be noticed that the observations, expressed in millions of euros, related to Credibility ($m = 13.33$, $SD = 29.47$) and Possibility ($m = 26.19$, $SD = 101.72$) present the highest variance. The distribution concerning Plausibility, instead, is characterized by the lowest variance ($m = 6.62$, $SD = 10$). The Plausibility opinions, which in the present study concern the decisions of experts, tend to occupy a definite area of hypotheses characterized by less uncertainty. As the considered field of opinions expands itself towards the areas of sentiment and other less relevant sources of information, the uncertainty on the players' evaluation increases, as the opinions are more heterogeneous.

Regarding the other features considered, the set concerning skills, age ($m = 24$, $SD = 3.66$, y.o.), position (e.g., offensive guard, wide receiver, etc.), wage ($m = 14.53$, $SD = 36.18$, millions of euros), height ($m = 182.48$, $SD = 7.13$, cm), weight ($m = 76.81$,

$SD = 7.74$, kg), preferred foot (right foot, left foot), and preferred positions was adopted. Non-numeric features, such as position or preferred foot, have been enumerated.

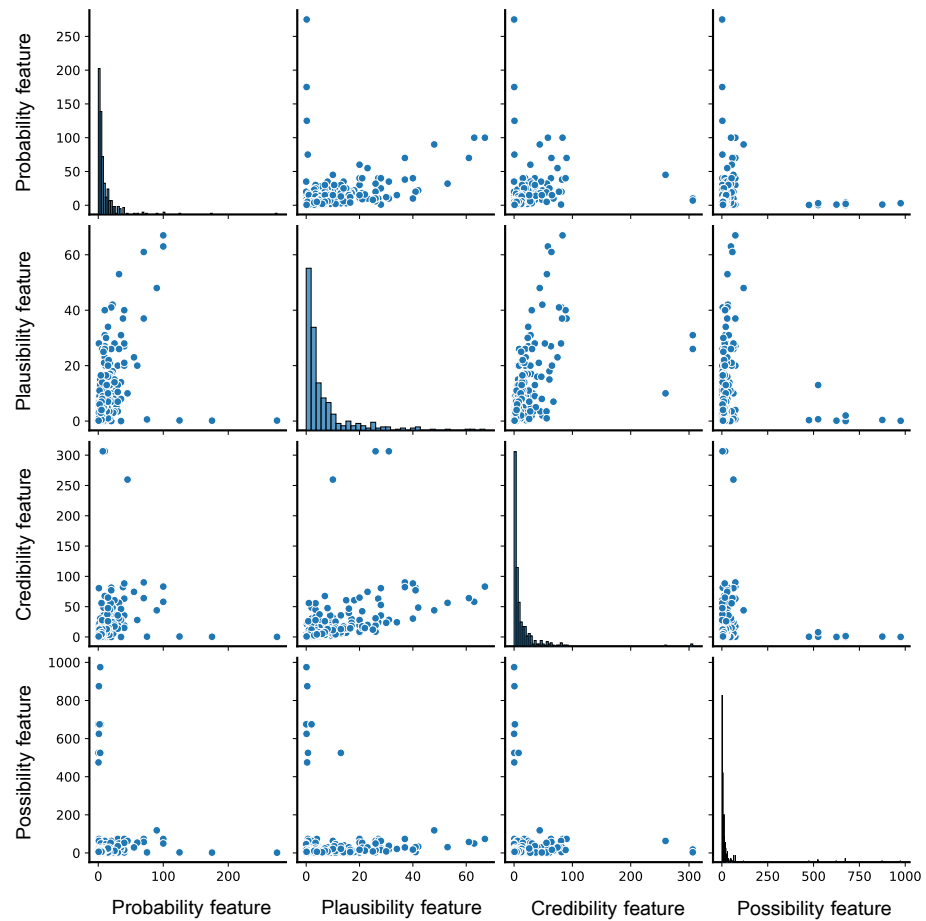


Figure 2. Pair distributions of features related to Probability, Plausibility, Credibility, and Possibility opinions.

In Figure 3, the box and whiskers plots related to the distributions of skills are shown. As organized in FIFA21, the scores for skills are specified as integer numbers in the range [0, 100].

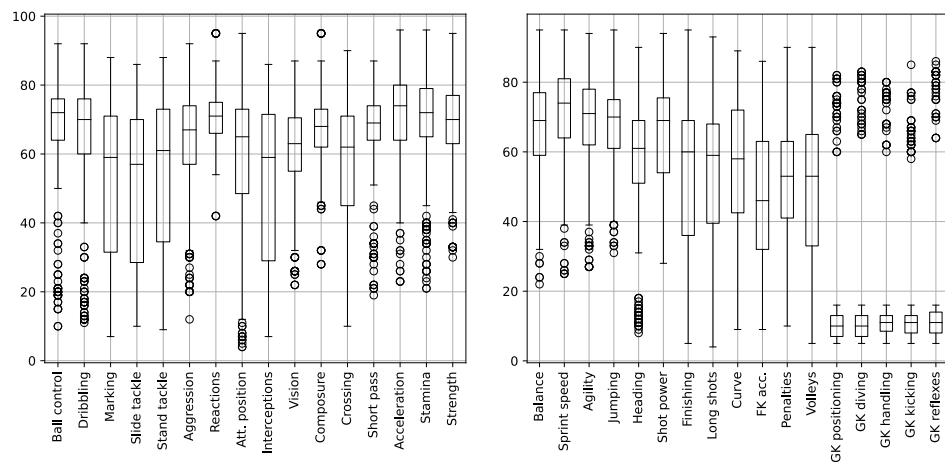


Figure 3. Distributions of features related to players' skills.

The attributes which present the highest dispersion are *making*, *slide tackle*, and *interceptions*, while those characterized by the highest amount of outliers are the dimensions reserved for goalkeepers, i.e., *GK positioning*, *GK diving*, *GK handling*, *GK kicking*, and *GK reflexes*. These last attributes present such characteristics due to the limited occurrence of goalkeepers in the dataset.

5.3. Implementation Details

The input features were pre-processed through a re-scaling in the $[-1, 1]$ range. For each input data point reserved for the backpropagation process, the related non-re-scaled values of Probability, Plausibility, Credibility, and Possibility are used by the L_{MBP} loss function to compute the prediction error. Thus, the network acquires two inputs: the first is the vector of re-scaled attributes and opinions, employed for performing the prediction; the second is the vector of non-re-scaled opinions, employed for computing the loss function. In our experiments, the network consists of three dense layers composed of 512, 256, and 16 neurons, respectively, all characterized by ReLu activation; the end-point, as already discussed in Section 4, consists of a four-neuron softmax layer.

The learning parameters of the network were initialized by sampling from a random uniform distribution.

5.4. Feature Selection

Feature selection was performed by computing feature importance through a brute force approach, as the number of features is limited. The employed process is described through the pseudo-code described in Algorithm 1, in which the *positiveImportanceScores* represents the list of importance scores concerning the considered set of features.

Algorithm 1 The employed algorithm for computing feature importance

```

importanceScores  $\leftarrow$  array[ $N_{features}$ ]
baseScore  $\leftarrow$  evaluate(compiledModel,  $X_{train}$ ,  $X_{test}$ ,  $y_{train}$ ,  $y_{test}$ )
index  $\leftarrow$  0
while index <  $N_{features}$  do
     $X_{train}^{new}$   $\leftarrow$  deleteFeatureAtIndex( $X_{train}$ , index)
     $X_{test}^{new}$   $\leftarrow$  deleteFeatureAtIndex( $X_{test}$ , index)
    predScore  $\leftarrow$  evaluate(compiledModel,  $X_{train}^{new}$ ,  $X_{test}^{new}$ ,  $y_{train}$ ,  $y_{test}$ )
    importanceScores[index]  $\leftarrow$  predScore – baseScore
end while
lowerBound  $\leftarrow$  –min(importanceScores)
positiveImportanceScores  $\leftarrow$  array[ $N_{features}$ ]
while index <  $N_{features}$  do
    positiveImportanceScores[index]  $\leftarrow$  lowerBound + importanceScores[index]
end while

```

The process starts by instantiating an array of null values characterized by $N_{features}$ dimensions, i.e., the number of considered dimensions, and by training and evaluating the compiled model on the original set of features. For each input dimension in the original set, a new set of inputs is generated by deleting the given feature associated with the considered dimension; then, the same model is trained and evaluated on the new input. At each iteration, the array initialized at the beginning is populated with the differences between the prediction scores obtained by evaluating the model with the original set of features and the temporary subsets of input dimensions. Finally, the result is obtained by adding the negation of the minimum occurrence to the elements of the array.

The procedure allows one to evaluate the importance of a given feature by computing the effect on the performances in terms of Mean Best Price Error. In the case in which the elimination of a given feature results in better performances, the related dimension is considered relevant for the decision. Its relevance is directly proportional to the difference between the errors obtained by evaluating the model with and without the given dimension.

5.5. Results

The best model obtained is characterized by online learning for backpropagation and Adam optimization at the learning rate of 10^{-2} . The proposed model was tested by considering different subsets of opinions to study the performances in different conditions of uncertainty. The experimentation was performed by considering the most important features obtained through the process described in Section 5.4.

In Table 1, the importance of the considered input dimensions is shown in the form of a ranking. In particular, the features providing an importance score lower than or equal to 0.056 contributed to a higher MBPE in the prediction.

Table 1. Results of the feature importance analysis obtained through the proposed method.

Feature	Importance	Feature	Importance
Composure	0.137	Long shots	0.073
GK reflexes	0.136	Short pass	0.072
Interceptions	0.13	Balance	0.071
Curve	0.129	Position	0.064
Acceleration	0.124	FM sentiment-society evaluation (P_3)	0.06
FIFA evaluation (P_4)	0.124	Vision	0.058
Stamina	0.102	GK handling	0.056
GK positioning	0.102	Finishing	0.05
Weight	0.1	Aggression	0.048
Volleys	0.099	FK accuracy	0.047
Marking	0.091	Crossing	0.037
Ball control	0.09	Penalties	0.037
Preferred positions	0.08	Slide tackle	0.034
Stand tackle	0.086	Heading	0.031
TM evaluation (P_1)	0.083	Dribbling	0.027
GK diving	0.078	Age	0.025
Strength	0.077	Jumping	0.023
FM value (P_2)	0.077	Agility	0.016
Reactions	0.076	GK kicking	0.014
Wage	0.075	Sprint speed	0.006
Height	0.074	Preferred foot	0.0

Table 2 shows the Mean Best Price Error, expressed in millions of euros, regarding different sets of opinions after feature selection. Except for the evaluation of the model by considering Plausibility, Credibility, and Possibility opinions individually, a significant increase in the performances was found after feature selection. However, the set of dimensions related to Credibility and Possibility increases the uncertainty in the prediction. The analysis of the differences between the groups of re-scaled input features associated with a prevalence, during the prediction phase, in one of the weights α_1 , α_2 , α_3 , and α_4 , provided significance for the dimensions of *height* ($p = 0.006$) and *marking* ($p = 0.039$). For all the experiments, the training and testing were performed on 338 and 60 samples, respectively.

The best performances were obtained by considering the set composed of Probability and Plausibility, as well as the Probability considered individually. The results identify substantial differences between the decision tests conducted using the different sets of opinions. For the problem addressed, concerning the prediction of the transfer cost of the

players, we obtained that the smallest error, in the prediction phase, is reached by using the decision associated with the Probability (0.86 MBPE, 0.61% of normalized error) only. The worst performances, however, were found individually considering the opinions of Plausibility, Credibility, and Possibility. This result corresponds to the prodromal theory of decision and reasoning in uncertainty and info-incompleteness conditions [1], since the obtaining, and the use, of particularly relevant opinions concerning the sphere of Probability represents a condition of info-completeness. Conversely, by eliminating the direct evidence, i.e., by neglecting the Probability opinion, there is a larger error in the prediction phase. The decision in conditions of info-incompleteness can introduce greater uncertainty in the decision phase, as the lack of direct evidence forces the decision-maker to evaluate the opinions deriving from experts, sentiments, and subjects of weaker relevance. The prediction problem addressed in the present study can be traced back to the hierarchical characterization of the *overlap model with a shift based on probability* taken up in Section 2. The decision hierarchy is in alignment with the results based on the increase in the error as a function of the type of opinion evaluated; for the present problem, the order of priority reflects the increasing order of Mean Best Price Error on the test set: (i) Probability (0.86 MBPE, 0.61% of normalized error); (ii) Credibility (2.34 MBPE, 1.79% of normalized error); (iii) Plausibility (2.48 MBPE, 0.90% of normalized error); (iv) Possibility (2.87 MBPE, 2.21% of normalized error).

Table 2. Performance in predicting the best price by considering different sets of opinions and the most important features.

Model Evaluation Considering the Most Important Features		
Chosen Opinions	Mean Best Price Error (mln)	Normalized Error (%)
Probability, Plausibility, Credibility, Possibility	1.01	0.72
Plausibility, Credibility, Possibility	2.25	1.72
Probability, Plausibility	0.91	0.64
Credibility, Possibility	2.41	1.85
Probability	0.86	0.61
Plausibility	2.48	1.90
Credibility	2.34	1.79
Possibility	2.87	2.21

Following the logic in the approach of *overlap with shift based on probability*, for each player, the most promising final decision is taken based on the following steps:

1. The final decision is made based on the Probability opinion, if available, with an expected error equal to 0.86 MBPE (0.61% of normalized error), proportional to the weight α_1 estimated by the model;
2. If the Probability opinion is not available, the final decision is made based on the Credibility opinion, if available, with an expected error of 2.34 MBPE (1.79% of normalized error), in proportion to the α_3 weight estimated by the model;
3. If the Probability and Credibility opinions are not available, the final decision is taken based on the Plausibility opinion, if available, with an expected error of 2.48 MBPE (0.90% of normalized error), in proportion to the α_2 weight estimated by the model;
4. If the Probability, Credibility, and Plausibility opinions are not available, the final decision is taken based on the Possibility opinion, if available, with an expected error of 2.87 MBPE (2.21% of normalized error), in proportion to the α_4 weight estimated by the model.

To optimize the performance in conditions of info-incompleteness, an extension can be added to the aforementioned steps that involve the simultaneous contribution of multiple opinions. In particular, in the case the opinions of Plausibility, Credibility, and Possibility are available, it is advisable to make the decision considering the weighted sum of the related contributions, as this solution provides a lower Mean Best Price Error than what would be obtained by considering only the opinion concerning Credibility.

By extending the logic of the *overlap with shift based on probability* approach to the joint evaluation of several opinions, the most promising final decision is optimized based on the following steps:

1. The final decision is made based on the Probability opinion, if available, with an expected error equal to 0.86 MBPE (0.61% of normalized error), proportional to the weight α_1 estimated by the model;
2. If the Probability opinion is not available, the final decision is made based on the opinions of Plausibility, Credibility, and Possibility, if available, with an expected error equal to 2.25 MBPE (1.72% of normalized error), based on the sum of the opinions weighted by the values α_2 , α_3 , and α_4 identified by the model;
3. If the Probability and Plausibility opinions are not available, the final decision is made based on the Credibility opinion, if available, with an expected error of 2.48 MBPE (1.90% of normalized error), in proportion to the α_3 weight estimated by the model;
4. If the Probability, Plausibility, and Credibility opinions are not available, the final decision is made based on the Possibility opinion, if available, with an expected error of 2.87 MBPE (2.21% of normalized error), in proportion to the α_4 weight estimated by the model.

It is further interesting to note that, unlike the starting theoretical model, in which Plausibility was characterized by a higher priority than Credibility, in this case, the opposite is true. For the problem of football players' evaluation, the experts' opinion is weaker.

To provide a visual example of the prediction performances, Figure 4 shows a comparison between the ground truth and the best price predicted by the model trained with the subset of the most important features (1.01 MBPE, 0.72% of normalized error).

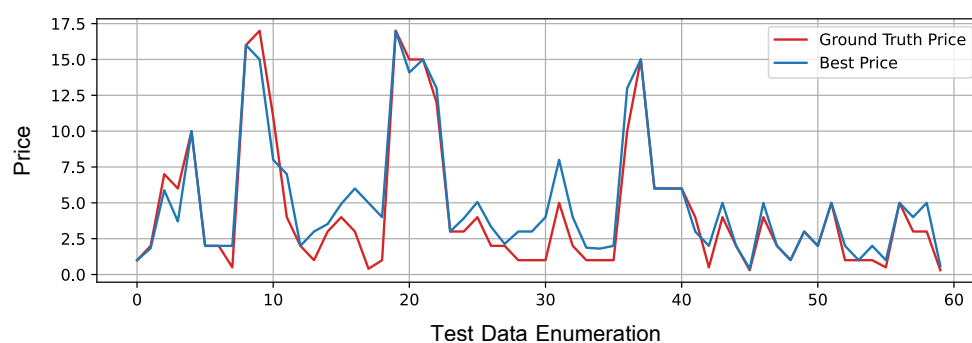


Figure 4. Comparison between the ground truth and the best price prediction obtained through the proposed model, after feature selection, based on Probability, Plausibility, Credibility, and Possibility opinions.

To validate the statistical evidence of the above results, it was decided to perform hypothesis testing concerning the differences between the models trained through different sets of opinions. To compute p-values, two-thousand bootstrapping sets have been generated from the predictions obtained by evaluating the models on the test set. Each generated set is compared with the ground truth to compute an MBPE score, which in turn is subtracted with the MBPE score obtained by evaluating another model. For each pair of models, a distribution of the differences is generated and the p-value is computed. Table 3 shows the results of the analysis by considering 0.05 as the threshold for significance.

The results show that, for the problem addressed in the present study, there is no statistical significance in adopting all the four opinions instead of considering the Proba-

bility only; to obtain optimal performances, there is strong evidence for considering the Probability, both exclusively and in conjunction with Plausibility, Credibility, and Possibility. Furthermore, the choice of considering both Credibility and Possibility is significantly better than that of considering Possibility only. Regarding other couples of hypotheses, no further statistical significance was found.

Table 3. Significance testing results regarding the differences in predicting the best price between pairs of opinions’ subsets.

p-Values of the Pair Differences								
	P_1, P_2, P_3, P_4	P_2, P_3, P_4	P_1, P_2	P_3, P_4	P_1	P_2	P_3	P_4
P_1, P_2, P_3, P_4	>0.05	0.002	0.29	<0.001	0.08	<0.001	<0.001	<0.001
P_2, P_3, P_4	0.002	>0.05	<0.001	0.512	<0.001	0.1	0.67	0.067
P_1, P_2	0.29	<0.001	>0.05	<0.001	0.620	<0.001	<0.001	<0.001
P_3, P_4	<0.001	0.512	<0.001	>0.05	<0.001	0.8	1.204	0.04
P_1	0.08	<0.001	0.620	<0.001	>0.05	<0.001	<0.001	<0.001
P_2	<0.001	0.1	<0.001	0.8	<0.001	>0.05	0.50	0.23
P_3	<0.001	0.67	<0.001	1.204	<0.001	0.50	>0.05	0.109
P_4	<0.001	0.067	<0.001	0.04	<0.001	0.23	0.109	>0.05

6. Discussion

The results obtained through the experimentation of the proposed model applied to the prediction of evaluation of football players show that the opinions related to Plausibility, Credibility, and Possibility are not useful in conditions of info-completeness. Strong evidence was found in employing the Probability only to make the final decision. This result is coherent with [10,11], as the crowdsourcing, from the introduction of Transfermarkt, has become the main source for evaluation of football players. Conversely, the opinions concerning Plausibility, Credibility, and Possibility are essential in conditions of info-incompleteness; in this context, the optimal decision can be achieved by adopting an arbitrary set of opinions, except for the couple Credibility–Possibility, which was found to be more promising than the Possibility considered exclusively.

In conclusion, the performance in estimating values of football players is 0.86 MBPE (0.61% of normalized error) in info-completeness conditions, while it is 2.47 ± 0.188 MBPE ($1.89 \pm 0.15\%$ of normalized error) in info-incompleteness.

The above results allow the research concerning Artificial Intelligence methodologies and decision support systems to be extended, as they provide an approach that potentially improves the hypotheses for the solution of predictive tasks. The approach tested in this study is simple and applicable to any decision context: if, in the prediction phase, one or more opinions regarding the output are available, it could be useful to consider these inputs to improve the performance of the model. For instance, in the case we want to consider, accuracy could be improved by a case study different from the one chosen in this work, such as the classification of images, the use of probable, plausible, credible, and possible opinions, if available. The conditional was used as the aforementioned field of application has yet to be tested; the results of this experiment could be different from those obtained in the present study (e.g., the Credibility opinion could provide better performances than the Probability).

The model proposed in the present study can be conceived as a sort of *human-in-the-loop* model, i.e., a predictor that requires human interaction. Instead of a human, our model proposes the Probability, Plausibility, Credibility, and Possibility opinions, provided by certain sets of humans, which can be extracted automatically by certain information sources. Having a human being available for supporting the model in real-time inference is very cost-effective; instead, in the case the same support is found in accessible information sources, such as the web, the external support to the model becomes cheaper. The potential

for improvement is high, as human-in-the-loop models were recently proved to be effective, especially in medicine and cybersecurity [22].

7. Conclusions and Future Work

In the present study, the problem of assessing the best opinion in conditions of uncertainty and info-incompleteness was addressed. To achieve this objective, we proposed a solution that provides a learning model that, starting from the observations related to Probability, Plausibility, Credibility, and Possibility, together with other relevant characteristics, provides the weights associated with the considered set of opinions. The proposed model minimizes the error of the results provided by the best price function, defined as the weighted sum of the considered opinions. The experiment was performed on a real case study concerning the market evaluation of soccer players by building a dataset based on the information sources of Transfermarkt, Football Manager, and FIFA21. The input space concerns features acquired one year before the subsequent transfer, while the ground truth is represented by the cost of the actual transfer. The experiments were carried out for a total of 398 occurrences by varying the set of opinions acquired for taking decisions, both in conditions of info-completeness and info-incompleteness; in the first case, the Probability was considered, while, in the second case, we limited the hypotheses to the set of Plausibility, Credibility, and Possibility only.

The results prove the consistency to the prodromal study taken as reference and that it is possible to reach an error for the price prediction of 0.86 MBPE (0.61% of normalized error) and 2.47 ± 0.188 MBPE ($1.89 \pm 0.15\%$ of normalized error) on the test set in conditions of info-completeness and info-incompleteness, respectively. Furthermore, from the analysis of statistical significance, it was found that the Probability opinion is fundamental in conditions of info-completeness; instead, in conditions of info-incompleteness, it is possible to adopt any set that considers Plausibility, Credibility, and Possibility. Finally, it was found that the employment of the Credibility–Possibility pair represents a better choice compared to the assumption involving the Possibility opinion only.

A possible future work regards the extension of the present study for the assessment of opinions' relevance by minimizing the occurrence and reliability functions in conditions of uncertainty and info-incompleteness. Regarding a real-world application, it is possible to define a decision support system to assist the Atmosphere Arc model [23], which brings the real economy into the digital economy through a Decentralized Content Management System (DCMS) and an Oracle. The DCMS contains the documentation of human work, while the Oracle analyzes the documentation and distributes a blockchain token. The proposed model can be employed to support token production and the documentation of the activities concerning a soccer society, providing a better estimation of the value.

Author Contributions: All authors contributed equally to the present work. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

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

References

1. Iovane, G.; Gironimo, P.D.; Chinnici, M.; Rapuano, A. Decision and Reasoning in Incompleteness or Uncertainty Conditions. *IEEE Access* **2020**, *8*, 115109–115122. [[CrossRef](#)]
2. Polya, G. *Mathematics and Plausible Reasoning: Induction and Analogy in Mathematics*; Princeton University Press: Princeton, NJ, USA, 1954.
3. Polya, G. *Mathematics and Plausible Reasoning: Patterns of Plausible Inference*; Princeton University Press: Princeton, NJ, USA, 1990.
4. Shafer, G. *A Mathematical Theory of Evidence*; Princeton University Press: Princeton, NJ, USA, 1976.

5. Dempster, A.P. Upper and Lower Probabilities Induced by a Multivalued Mapping. *Ann. Math. Stat.* **1967**, *38*, 325–339. [[CrossRef](#)]
6. Smarandache, F.; Dezert, J. An introduction to DSMT. In *Advances and Applications of DSMT for Information Fusion (Collected Works)*; American Research Press (ARP): Rehoboth, NM, USA, 2009; Volume 3.
7. Dobson, S.; Goddard, J.A.; Dobson, S. *The Economics of Football*; Cambridge University Press: Cambridge, UK, 2017.
8. Behravan, I.; Mohammad, M.R.S. A novel machine learning method for estimating football players' value in the transfer market. *Soft Comput.* **2021**, *25*, 2499–2511. [[CrossRef](#)]
9. Müller, O.; Simons, A.; Weinmann, M. Beyond Crowd Judgments: Data-Driven Estimation of Market Value in Association Football. *Eur. J. Oper. Res.* **2017**, *263*, 611–624. [[CrossRef](#)]
10. Singh, P.; Lamba, P. Influence of crowdsourcing, popularity and previous year statistics in market value estimation of football players. *J. Discret. Math. Sci. Cryptogr.* **2019**, *22*, 113–126. [[CrossRef](#)]
11. Steffen, H.; Callsen-Bracker, H.; Kreis, H. When the crowd evaluates soccer players' market values: Accuracy and evaluation attributes of an online community. *Sport Manag. Rev.* **2014**, *17*, 484–492.
12. Pérez-González, B.; Fernández-Luna, A.; Vega, P.; Burillo, P. The relative age effect: Does it also affect perceived market value? The case of the Spanish LFP (Professional Football League). *J. Phys. Educ. Sport* **2018**, *18*, 1408–1411.
13. Majewski, S. Identification of Factors Determining Market Value of the Most Valuable Football Players. *Cent. Eur. Manag. J.* **2016**, *24*, 91–104. [[CrossRef](#)]
14. Yigit, A.T.; Samak, B.; Kaya, T. Football player value assessment using machine learning techniques. In *International Conference on Intelligent and Fuzzy Systems*; Springer: Cham, Switzerland, 2019; pp. 289–297.
15. Felipe, J.L.; Fernandez-Luna, A.; Burillo, P.; Riva, L.E.D.L.; Sanchez, J.; Garcia-Unanue, J. Money talks: Team variables and player positions that most influence the market value of professional male footballers in Europe. *Sustainability* **2020**, *12*, 3709. [[CrossRef](#)]
16. Franck, E.; Nüesch, S. Talent and/or popularity: What does it take to be a superstar? *Econ. Inq.* **2012**, *50*, 202–216. [[CrossRef](#)]
17. Kiefer, S. The Impact of the Euro 2012 on Popularity and Market Value of Football Players. *Int. J. Sport Financ.* **2014**, *9*, 95–110.
18. Yaldo, L.; Shamir, L. Computational Estimation of Football Player Wages. *Int. J. Comput. Sci. Sport* **2017**, *16*, 18–38. [[CrossRef](#)]
19. Devlin, J.; Chang, M.; Lee, K.; Toutanova, K. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *arXiv* **2019**, arXiv:1810.04805.
20. GitHub. IBM Developer Model Asset Exchange: Text Sentiment Classifier. Available online: <https://github.com/IBM/MAX-Text-Sentiment-Classifier> (accessed on 23 November 2021).
21. IBM Research. IBM Project Debater. Available online: https://research.ibm.com/haifa/dept/vst/debating_data.shtml (accessed on 23 November 2021).
22. Zagalsky, A.; Te'eni, D.; Yahav, I.; Schwartz, D.G.; Silverman, G.; Cohen, D.; Mann, Y.; Lewinsky, D. The design of reciprocal learning between human and artificial intelligence. *Proc. ACM-Hum.-Comput. Interact.* **2021**, *5*, 443. [[CrossRef](#)]
23. Atmosphere Arc Whitepaper (2018). Available online: <https://atmospherearc.com/atmospherearc.pdf> (accessed on 24 November 2021).