# Fuzzy Naive Bayesian Classification in RoboSoccer 3D: A Hybrid Approach to Decision Making

Carlos Bustamante, Leonardo Garrido, and Rogelio Soto

Center for Intelligent Systems
Monterrey Institute of Technology
Monterrey NL 64849, Mexico
{cfbh,leonardo.garrido,rsoto}@itesm.mx

**Abstract.** We propose the use of a Fuzzy Naive Bayes classifier with a MAP rule as a decision making module for the RoboCup Soccer Simulation 3D domain. The Naive Bayes classifier has proven to be effective in a wide range of applications, in spite of the fact that the conditional independence assumption is not met in most cases. In the Naive Bayes classifier, each variable has a finite number of values, but in the RoboCup domain, we must deal with continuous variables. To overcome this issue, we use a fuzzy extension known as the Fuzzy Naive Bayes classifier that generalizes the meaning of an attribute so it does not have exactly one value, but a set of values to a certain degree of truth. We implemented this classifier in a 3D team so an agent could obtain the probabilities of success of the possible action courses given a situation in the field and decide the best action to execute. Specifically, we use the pass evaluation skill as a test bed. The classifier is trained in a scenario where there is one passer, one teammate and one opponent that tries to intercept the ball. We show the performance of the classifier in a test scenario with four opponents and three teammates. After a brief introduction, we present the specific characteristics of our training and test scenarios. Finally, results of our experiments are shown.

## 1 Introduction

Classification is a statistical operation in which certain objects are put into groups or classes according to their characteristics, sometimes called attributes, found on a training set. There are many approaches to classification in literature, like Decision trees, Neural networks, Support vector machines and Bayesian networks, among others. From the aforementioned classifing methods, the bayesian approach is the most commonly used to deal with uncertainty, because it is based on the probability theory.

A well known classifier is the Naive Bayes classifier [1], a simple type of bayesian network [2] that explodes the conditional independence assumption among attributes given the class. In real life, this assumption does not hold most of the time. However, Naive Bayes classifiers have proven to be successful

G. Lakemeyer et al. (Eds.): RoboCup 2006, LNAI 4434, pp. 507–515, 2007.

and more or equally effective than other classification methods in certain domains, like anti-spam filtering [3], information retrieval [4], speech recognition [5], emotion recognition [6] and medicine [7]. An analysis of the reasons why Naive Bayes works well is done in [8].

Generally, in a Naive Bayes classifier the attributes are discrete, but in most real-life situations, attributes are continuous. Crisp partitioning the domain can clearly cause some loss of information [9]. There have been some approaches to overcome this matter with the use of fuzzy variables [10,11].

The aim of this paper is to use the Fuzzy Naive Bayes classifier proposed in [10] as a decision module for a RoboCup [12] Soccer Simulation 3D team. Although the RoboCup simulation 3D is based on SPADES [13] multi-agent discrete event simulator, the data handled by the soccer agents is continuous, i.e. is defined in the range of real numbers. Also, the environment has too many features and the sensed data is influenced by a random noise. These characteristics make the RoboCup simulation 3D domain an excelent platform to test the Fuzzy Naive Bayes classifier effectiveness.

## 2 Fuzzy Naive Bayes Classifier

The Naive Bayes probabilistic model is one of the simplier Bayesian Network models used in Artificial Intelligence and Machine Learning nowadays. Let $C$ be a class label with $k$ possible values, and $X_1...X_n$ be a set of attributes or features of the world with a finite domain $D(X_i)$ where $i = 1..n$. The objective is to obtain the conditional model $P(C|X_1, ..., X_n)$. We can represent this model using the bayes' rule as follows

$$P(C|X_1, ..., X_n) = \frac{P(X_1, ..., X_n|C)P(C)}{P(X_1, ..., X_n)} \tag{1}$$

where $P(C)$ and $P(X_1, ..., X_n)$ are a priori probabilities, and $P(X_1, ..., X_n|C)$ is the likelihood of event $X_1, ..., X_n$ conditioned on the class $C$. Notice that the denominator remains constant for every value of $C$, thus it serves as a normalization constant and will be ommited from now on.

For computing the conditional probabilities of the model, the full joint probability table is needed. When the number of attributes is very large or the domain of such attributes consists of a large set of values, the use of the full joint becomes unfeasible. To overcome this difficulty, the conditional independence assumption is exploded. If we assume that every attribute is independent of each other given the class $C$, the model can be reformulated again as

$$P(C|X_1, ..., X_n) = P(C) \prod_{i=1}^{n} P(X_i|C) \tag{2}$$

that is known as the *Naive Bayes probabilistic model.* because the assumption of conditional independence does not hold in most scenarios. Besides of that, the Naive Bayes model has proven to be effective in a whole range of applications.

The *Naive Bayes Classifier* combines the Naive Bayes probabilistic model with a so called decision rule (or discriminant function in [8]). Generally, a *maximum a posteriori (MAP)* decision rule is used and we get the definition of a *Naive Bayes Classifier*

$$NBclassify(a) = \arg\max_{c \in C} P(c) \prod_{i=1}^{n} P(x_i|c) \tag{3}$$

where $x_i$ means $X_i = x_i$, $c$ means $C = c$ and $a$ is a complete assignation of attributes, i.e. $a = \{X_1 = x_1, ..., X_n = x_n\}$ and so will be hereafter. In this context, $a$ represents a new example not classified yet. The classifier is used to select a class $C$ given the new example $a$, based on the previously calculated values of all the probabilities needed by the model. The probabilities are estimated using relative frequencies from data. Sometimes, Laplace-correction is applied to smooth calculations avoiding extreme values obtained with small training sets.

The classical Naive Bayes classifier considers attributes and classes with discrete domains. When dealing with continuous domains, the classifier needs a modification. One way is discretizing or crisp partitioning the domain of attributes into a finite number of classical sets. But that could cause a loss of information [9].

A better method is proposed in [10], consisting of a hybrid classifier bringing together Fuzzy Set Theory and a Naive Bayes classifier, named the *Fuzzy Naive Bayes classifier*

$$FNBclassify(a) = \arg\max_{c \in C} P(c) \sum_{x_1 \in X_1} P(x_1|c)\mu_{x_1} \; ... \sum_{x_n \in X_n} P(x_n|c)\mu_{x_n} \tag{4}$$

where $\mu_{x_i} \in [0,1]$ denotes a membership fuction or degree of truth of attribute $x_i \in X_i$ in a new example $a$. To be conservative, it is required that all degrees of truth are normalized in the current variable assignation, in this case $\sum_{x_i \in X_i} \mu_{x_i} = 1$. The probabilities for equation (4) can be calculated as below

$$P(C = c) = \frac{(\sum_{e \in L} \mu_c^e) + 1}{|L| + |D(C)|} \tag{5}$$

$$P(X_i = x_i) = \frac{(\sum_{e \in L} \mu_{x_i}^e) + 1}{|L| + |D(X_i)|} \tag{6}$$

$$P(X_i = x_i | C = c) = \frac{(\sum_{e \in L} \mu_{x_i}^e \mu_c^e) + 1}{(\sum_{e \in L} \mu_c^e) + |D(X_i)|} \tag{7}$$

where $L$ is the training set consisting of all examples $e = \{X_1 = x_1, ..., X_n = x_n, C = c\}$, $\mu_c^e \in [0,1]$ denotes the degree of truth of $c \in C$ in a example $e \in L$, and $\mu_{x_i}^e \in [0,1]$ is the membership of attribute $x_i \in X_i$ in such example. As mentioned before, all degrees of truth must be normalized such that $\sum_{c \in C} \mu_c^e = 1$ and $\sum_{x_i \in X_i} \mu_{x_i}^e = 1$. Notice that Laplace-correction is applied to compute the probabilities.

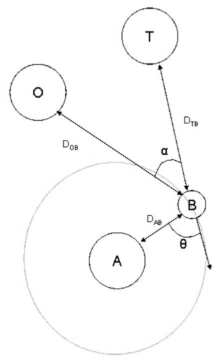


**Fig. 1.** Training scenario used to learn the probabilities for the Fuzzy Naive Bayes classifier, where we can see a passer agent (A), a receiver teammate (T), an opponent (O) and the ball (B). The features taken into account are: distance to ball $d_{AB}$, distance to teammate $d_{TB}$, distance to opponent $d_{OB}$, alignment angle ($\theta$) and angle bewteen teammate and opponent from the ball's view point ($\alpha$).

## 3 Empirical Scenarios

There are many opportunities to apply and test the classifier in the RoboCup domain. We chose the pass skill as a platform because it is a classical test bed for RoboCup. Specifically, we focus on pass evaluation, i.e., the capacity of an agent to evaluate the probability of success of a pass.

We created a training scenario for learning the probabilities for the Fuzzy Naive Bayes model. It is similar to the scenario proposed in [14] for the 2D league. But here we have to consider the alignment angle and ball distance, because in 3D soccer the agents are spheres and sometimes have to surround the ball for kicking in the right direction. The training scenario is shown in figure 1.

The scenario is mounted as follows

1. A passer agent is placed in the center of the field.
2. The ball is randomly placed next to the passer.
3. A teammate is randomly placed at a distance $d_{TB} \in [2, 20]$ from the ball.
4. An opponent is randomly placed at a distance $d_{OB} \in [2, 20]$ from the ball, such that the angle between the teammate and the opponent from the ball's view point is $\alpha \in [0, \frac{\pi}{6}]$.

Before an episode begins, the agent records five variables: distance to the ball $d_{AB}$, distance to teammate $d_{AT}$, distance to opponent $d_{AO}$, alignment angle $\theta \in [0, \pi]$ and the angle between teammate and opponent $\alpha$. During each episode, the agent aligns with the ball to pass it to the teammate. Then both the teammate and the opponent try to intercept the pass. If the teammate gets the ball the episode is labeled as *SUCCESS*. If the opponent gets the ball first the episode is labeled as *MISS*.

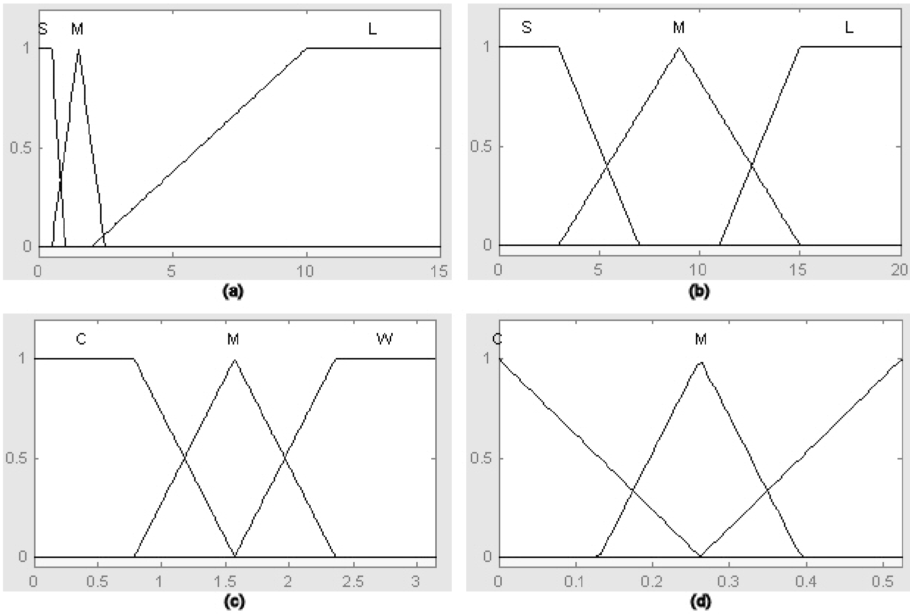


**Fig. 2.** Fuzzy Sets for each Fuzzy Variable. (a) Distance to the ball $d_{AB}$, (b) Distance to teammate $d_{AT}$ and distance to opponent $d_{AO}$, (c) Alignment Angle $\theta$ and (d) Angle between teammate and opponent $\alpha$.

Each variable mentioned above is a fuzzy variable and it is defined by several fuzzy sets. The fuzzy sets for distance to the ball $d_{AB}$, distance to teammate $d_{AT}$ and distance to opponent $d_{AO}$ variables are {*short, medium, long*}. The fuzzy sets for $\theta$ and $\alpha$ variables are {*closed, medium, wide*}. A graphical representation for each fuzzy variable is shown in figure 2.

## 4 Experimental Results

We obtained 1000 training examples, which we used to calculate the probabilities for the Fuzzy Naive Bayes classifier represented by equations (5), (6) and (7).

We ran 250 episodes to create a test set so we can measure the performance of the classifier. A representative graph is shown in figure 3. As can be seen in the figure, the graph stabilizes quickly, approximately at 120 traning examples. The maximum proportion of correclty classified examples was 0.806 which occured approximately at 350 examples.

To test the performance of the classifier for the pass evaluation skill, we created a test scenario which is shown in figure 4. In this scenario, the ball is in $(x = -20, y = 0)$ and the agent is placed at $(x \in [-22, -18], y \in [kickrange, 2])$. Four opponents and three teammates appear randomly in the area defined by $(x \in [-30, -10], y \in [10, 30])$.

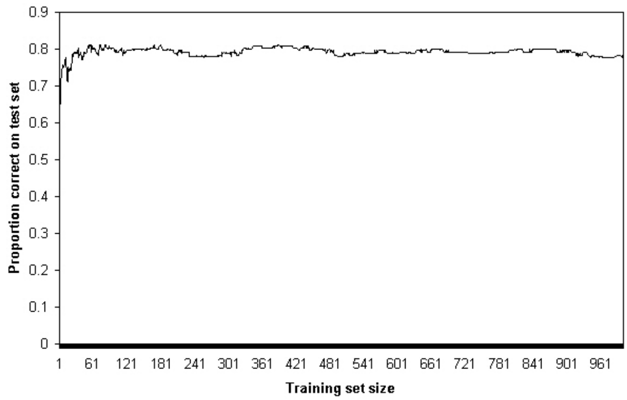


**Fig. 3.** Performance of the Fuzzy Naive Bayes classifier. The x-axis shows the number of examples used to train the classifier. The y-axis shows the proportion of the correcly classified examples in the test set. The size of the test set is 250.

When the scenario begins, the passer agent chooses the teammate with the best chances to intercept the ball using the Fuzzy Naive Bayes classifier. This is achieved in the following way: the passer uses the classifier to evaluate all 1 vs. 1 competitions between each teammate and each opponent. The lowest probability of success is stored for each teammate given all its 1 vs. 1 competitions and finally the teammate with the maximum probability of success is chosen. Formally,

$$Receiver = argmax_{t \in T}\ argmin_{o \in O}\ P(SUCCESS_{to}) \tag{8}$$

where T is the set of all teammates, O is the set of all opponents and $P(SUCCESS_{to})$ is the probability of success given the 1 vs. 1 competition between teammate $t \in T$ and opponent $o \in O$.

An episode is considered *SUCCESS* if a teammate is able to intercept the ball before any opponent does. If an opponent gets the ball before a teammate does, the episode is classified as *MISS*.

We obtained 300 examples using the test scenario described above. The percentage of SUCCESS and MISS classified examples is shown in table 1.

**Table 1.** Percentage of SUCCESS and MISS for a total of 300 classified examples using the test scenario of figure 4

| Class | Percentage | Number of examples |
|---|---|---|
| SUCCESS | 76 | 228 |
| MISS | 24 | 72 |

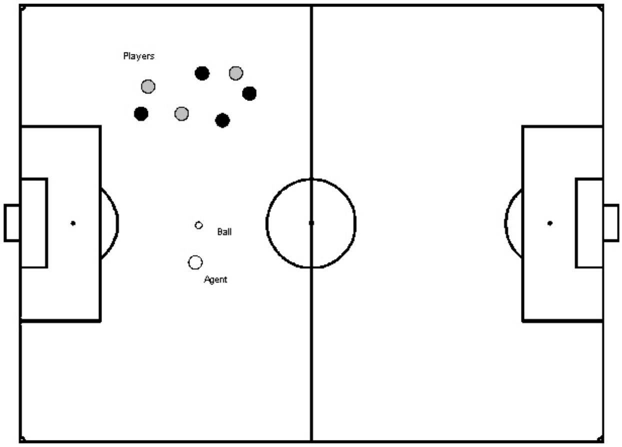


**Fig. 4.** Test scenario for the pass evaluation skill. Four opponent agents (black circles) and three teammates (gray-circles) appear randomly in a certain area of the field. The passer (white circle) and the ball (little circle) are placed a few meters away from them. The passer chooses the teammate with the best chances to intercept the ball using the classifier.

## 5 Conclusions

In this paper, we researched the application of a Fuzzy Naive Bayes classification algorithm to the decision making process of a RoboCup 3D team, specifically in the pass evaluation skill. The Naive Bayes method has proved to be effective in a wide range of situations although conditional independence assumtion is not met. As in RoboCup simulation we can consider variables as being continuous, we suggested seeing them as fuzzy variables so we could apply the Fuzzy extension to Naive Bayes proposed in [10].

We trained the classifier under a specific scenario where one agent passes the ball to a teammate and an opponent tries to intercept such pass. This 1 vs. 1 competition for pass evaluation can be easily extended to be used for the pass selection skill as in equation (8). As shown in figure 3 the classifier (trained with 1000 training examples) correclty classified a proportion of 0.806 examples of a test set of 250. Although the performance is not so good as in other implementations of Fuzzy Bayes in similar domains as in [15], we think it is not a bad performance at all. In table 1 we can see that 76% of the passes in our test scenario where successful passes. This is relatively better than the results shown by Stone [16], who used a Decision Tree for the pass evaluation procedure and a similar test scenario, and just 65% of all passes where successful.

We plan to extend our implementation in a near future to other skills like dribble and shoot. We think the performance may increase if the fuzzy sets are constructed more carefully. Perhaps it would be possible to extend this classifier to use fuzzy k-means clustering to obtain the fuzzy sets directly from data. We would like to implement other classifiers and compare their performance against the Fuzzy Naive Bayes classifier. Another possibility is to use the classifier to decide when to execute an air kick and take advantage of the characteristics of the 3D environment. Our final goal is to have a 3D team fully based on the fuzzy-bayes hybrid approach for the world cup competitions.

## Acknowledgements

This work was supported in part by the research grant CAT011 and the Center for Intelligent Systems at Monterrey Institute of Technology (ITESM), Mexico.

## References

1. Langley, P., Iba, W., Thompson, K.: An Analysis of Bayesian Classifiers. In: Proc. 10th Nat. Conf. on Artificial Intelligence, pp. 223–228. AAAI Press and MIT Press, Cambridge, USA (1992)
2. Heckerman, D.: A tutorial on learning with bayesian networks. Technical Report MSR-TR-95-06, Microsoft Research, Redmond, Washington (1995)
3. Androutsopoulos, I., Koutsias, J., Chandrinos, K.V., Paliouras, G., Spyropoulos, C.D.: An Evaluation of Naive Bayesian Anti-Spam Filtering. In: Proceedings of the workshop on Machine Learning in the New Information Age (2000)
4. Lewis, D.: Naive Bayes at forty: The independence assumption in information retrieval. In: Proceedings of European Conference on Machine Learning, pp. 4–15 (1998)
5. Tóth, L., Kocsor, A., Csirik, J.: On Naive Bayes in Speech Recognition. Int. J. Appl. Math. Comput. Sci. 15(2), 287–294 (2005)
6. Sebe, N., Cohen, I., Garg, A., Lew, M.S., Huang, T.S.: Emotion recognition using a Cauchy naive Bayes classifier. In: Proceedings of 16th International Conference on Pattern Recognition, pp. 17–20 (2002)
7. Demsar, J., Zupan, B., Kattan, M.W., Beck, J.R., Bratko, I.: Naive Bayesian-based nomogram for prediction of prostate cancer recurrence. Medical Informatics Europe '99, Studies in health technology and informatics 68, 436–441 (1999)
8. Rish, I.: An empirical study of the naive bayes classifier. In: Proceedings of IJCAI-01 workshop on Empirical Methods in AI, pp. 41–46 (2001)
9. Friedman, N., Goldszmidt, M.: Discretization of continuous attributes while learning Bayesian networks. In: Saitta, L. (ed.) Proceedings of 13-th International Conference on Machine Learning, pp. 157–165 (1996)
10. Störr, H.-P.: A compact fuzzy extension of the Naive Bayesian classification algorithm. In: Proceedings InTech/VJFuzzy, pp. 172–177 (2002)
11. Tang, Y., Pan, W., Li, H., Xu, Y.: Fuzzy Naive Bayes classifier based on Fuzzy Clustering. In: IEEE International Conference on Systems, Man and Cybernetics (2002)
12. Kitano, H., Asada, M., Kuniyoshi, Y., Noda, I., Osawa, E.: RoboCup: The robot world cup initiative. In: Proceedings of the First International Conference on Autonomous Agents, pp. 340–347 (1997)
13. Riley, P., Riley, G.: SPADES: - A Distributed Agent Simulation Environment with Software-in-the-Loop Execution. In: Winter Simulation Conference Proceedings, pp. 817–825 (2003)

14. Buck, S., Riedmiller, M.: Learning Situation Dependent Success Rates Of Action In A RoboCup Scenario. In: Pacific Rim International Conference on Artificial Intelligence, p. 809 (2000)
15. Mostafa, M.G.-H., Perkins, T.C., Farag, A.A.: A Two-Step Fuzzy-Bayesian Classification for High Dimensional Data. In: 15th International Conference on Pattern Recognition (ICPR'00), vol. 3, pp. 3421–3424 (2000)
16. Stone, P.: Layered Learning in Multiagent Systems: A Winning Approach to Robotic Soccer. MIT Press, Cambridge (2000)