GENERALIZED NET MODEL OF THE INTUITIONISTIC FUZZY VERSION OF A NEAREST PROTOTYPE CLASSIFICATION METHOD

Stefan T. Hadjitodorov and Krassimir T. Atanassov

Centre for Biomedical Engineering – Bulgarian Academy of Sciences,
Acad. G. Bonchev Str., Bl. 105, Sofia-1113, BULGARIA
e-mails: sthadj@argo.bas.bg, krat@bas.bg

Abstract:

An intuitionistic fuzzy version of one of the statistical nonparametrical methods, the Nearest Prototype (NP) method, has been proposed. The degrees of membership, nonmembership and indeterminacy (uncertainty) of a pattern with unknown classification to a given class have been evaluated taking into account the distance to the prototype (mean) vector of that class. The confidence in the classification made is enhanced by means of a procedure of moving of the pattern towards the classes. The moving is based on the correlation coefficients and the distances to the class prototypes. In this way the degrees are being modified and adjusted, getting more precise values and aiming to decrease and possibly eliminate the indeterminacy. The procedure stops when the uncertainty is driven negligibly small. A Generalized Net (GN – an extension of the Petri net) model of the process of realizing of the above method is described. The model is an element of the series of authors’ investigations on the application of the GNs to the problems of pattern recognition.

2000 Mathematics Subject Classification. Primary 68Q90; Secondary 92B20
Keywords and phrases. Fuzzy classification, Generalized net, Intuitionistic fuzzy set, Model, Pattern recognition
1 Introduction

The Nearest Prototype (NP) method is one of the most often applied and accurate methods in pattern recognition [2, 4, 5, 14, 16, 17]. This method is a simplification of the well-known K-nearest neighbors (K-NN) classification method. The basic assumption of this statistical and nonparametrical method is that inside of the small hypersphere around a given pattern the probability density function is approximately constant [7]. For that reason the classifiers have to be restricted to a small value of the radius of the hypersphere. However, in cases of small sample size and overlapping classes this requirement causes accuracy decreasing. Many of modifications of the method, which improve it with respect to increasing the classification accuracy and minimization of the calculations, have been published. Several fuzzy versions have also been developed [2, 4, 13, 7, 14, 15, 17]. In [17] it has made an attempt to combine the advantages of the fuzzy description of the pattern and classes. In that paper, the idea is to include expert experience and knowledge in pattern recognition in a new way. In this way the pattern distribution in the space is implicitly taken into consideration. The goal is to cope with cases of small sample size and overlapping classes. A transformation of the distances using fuzzy description (degree of membership) is proposed and it is proved that the probability of error in that case is less than or equal to the probability of error in case of classical K - NN method (with nontransformed distances). In [8] an improvement of this method by means of application of degree of nonmembership besides the degree of membership is described and later tested [9]. In [10], the idea to combine the statistical and fuzzy approaches in pattern recognition, which intuitively leads to improvement of the recognition results, is further developed. A procedure which adjusts the degree of membership to the class (makes it more precise) and decreases the degree of indeterminacy and in such a way enhances the confidence in the classification is proposed. The adjusting procedure is based on moving the pattern under classification towards the class prototype (the mean) vector, thus changing the value of the degrees of membership and nonmembership.

The Generalized Net (GN; [1]) is an extension of the Petri net. Here a GN-model of the process of realizing of the above method is described. The model is an element of the series of authors’ investigations on the application of the GNs to the pattern recognition problems.

2 Short remarks on generalized nets

The concept of a Generalized Net (GN) is described in [1]. Some GNs may not have some of the components, thus giving rise to a special class of GNs called “reduced
GNs”. For the needs of the present research we shall use (and describe) one of the reduced types of GNs.

Formally, every transition of this reduced class of GNs is described by (Fig. 1):

$$Z = (L', L'', r, \Box),$$

where:

(a) $L'$ and $L''$ are finite, non-empty sets of places (the transition’s input and output places, respectively). For the transition in Fig. 1 these are

$$L' = \{l'_1, l'_2, \ldots, l'_m\}$$

and

$$L'' = \{l''_1, l''_2, \ldots, l''_n\};$$

(b) $r$ is the transition’s condition determining which tokens will pass (or transfer) from the transition’s inputs to its outputs; it has the form of an Index Matrix (IM; see [1]):

$$r = \begin{array}{c|cccc}
& l''_1 & \ldots & l''_j & \ldots & l''_n \\
\hline
l'_1 & r_{i,j} \\
\vdots & \vdots \\
l'_{i'} & (r_{i,j} - \text{predicate}) \\
\vdots & (1 \leq i \leq m, 1 \leq j \leq n) \\
l'_{m'} & \end{array}$$

$r_{i,j}$ is the predicate which corresponds to the $i$-th input and $j$-th output places. When its truth value is “true”, a token from the $i$-th input place can be transferred to the $j$-th output place; otherwise, this is not possible;
(c) \( \square \) is a Boolean expression. It may contain as variables the symbols which serve as labels for transition's input places, and it is an expression built up from variables and the Boolean connectives \( \land \) and \( \lor \) whose semantics is defined as follows:

\[
\land(l_{i_1}, l_{i_2}, \ldots, l_{i_u}) - \text{every place } l_{i_1}, l_{i_2}, \ldots, l_{i_u} \text{ must contain at least one token,}
\]

\[
\lor(l_{i_1}, l_{i_2}, \ldots, l_{i_u}) - \text{there must be at least one token in all places } l_{i_1}, l_{i_2}, \ldots, l_{i_u},
\]

where \( \{l_{i_1}, l_{i_2}, \ldots, l_{i_u}\} \subset L' \).

When the value of a type (calculated as a Boolean expression) is “true”, the transition can become active, otherwise it cannot.

The ordered four-tuple

\[
E = \langle A, K, X, \Phi \rangle
\]

is called simplest reduced Generalized Net (briefly, we shall use again “GN”) if:

(a) \( A \) is a set of transitions;

(b) \( K \) is the set of the GN’s tokens.

(c) \( X \) is the set of all initial characteristics the tokens can receive when they enter the net;

(d) \( \Phi \) is a characteristic function which assigns new characteristics to every token when it transfers from an input to an output place of a given transition.

Over the GNs a lot of types of operators are defined. One of these types is the set of hierarchical operators. One of them changes a given GN-place with a whole subnet (see [1]). Below, having in mind this operator, we shall use three places that will represent three separate GNs, constructed by the authors early.

3 A generalized net model of the intuitionistic fuzzy version of a nearest prototype classification method

Below we shall construct a reduced GN (Fig. 2) without temporal components, without transitions’, places’ and tokens’ priorities and without places’ and arcs’ capacities. Tokens there will keep all their history.

We shall describe the transition condition predicates and the tokens characteristics not fully formally for the sake at easier understanding of the formalism in use.

Initially, token \( \alpha \) enters place \( l_1 \) with an initial characteristic:

“training (sample) set of patterns (vectors) \( X = \{x_1, x_2, \ldots, x_N\} \), where each

\[
\text{pattern } x_j \in \mathcal{R}^n \text{ is described by } n \text{ features, } x_j = \langle x_{j,1}, x_{j,2}, \ldots, x_{j,n} \rangle \\
(j = 1,2,\ldots,N); \text{ - } \omega = \{\omega_1, \omega_2, \ldots, \omega_M\} \text{ - a set of classes defined over } \mathcal{R}^n,\]
such that class $\omega_i$ contains $N_i$ patterns ($i = 1, 2, ..., M$) and

$$\sum_{i=1}^{M} N_i = N'',$$

where $\mathcal{R}$ is the set of the real numbers.

$$Z_1 = <\{l_1\}, \{l_2\}, \frac{l_2}{l_1} \text{true} \lor (l_1) >.$$

The tokens obtain the characteristic

“mean vector $\mu_i = \frac{1}{N_i} \sum_{l=1}^{N_i} x_l$ and the standard deviation of the classes

$$\sigma_{i,k} = \sqrt{\frac{1}{N_i - 1} \sum_{l=1}^{N_i} (x_{l,k} - \mu_{i,k})^2} \ (k = 1, 2, ..., n; i = 1, 2, ..., M)$$

in place $l_2$.

Fig. 2: A GN-model

Initially, token $\beta$ enters place $l_3$ with an initial characteristic:

“pattern with unknown classification $x_u$”

$$Z_2 = <\{l_2, l_3, l_9\}, \{l_4, l_5\}, \frac{l_4}{l_5} \text{true} \lor false \lor \lor (l_2, l_3, l_9) >,$$

$$\frac{l_2}{l_3} \text{true} \lor false \lor W_{3,5} \lor (l_2, l_3, l_9), \frac{l_4}{l_5} \text{true} \lor false \lor false \lor false \lor false \lor false \lor false \lor false.$$


where
\[ W_{3,5} = W_{12,11} \lor W_{14,16} \]
(the descriptions of predicates \( W_{12,11} \) and \( W_{14,16} \) are given below).

The token \( \alpha \) enters place \( l_4 \) (from place \( l_2 \) or \( l_{15} \)) and obtains the characteristic

“determination of the distances between \( x_u \) and class centers using formula

\[ d_i = \| x_u - \bar{x}_i \| = \left[ \sum_{j=1}^{n} (x_{u,j} - \bar{x}_{i,j})^2 \right]^{1/2} \]

\( i = 1, 2, \ldots, M \),”

where these distances are ordered in an ascending way

\[ D = \langle d_{\text{min}}, \ldots, d_{\text{max}} \rangle, \]

while token \( \beta \) enters place \( l_5 \) in the final moment of the GN-functioning with the final characteristic

“pattern classification of \( x_u \) to the respective class with maximal \( \mu \)-value”.

\[ Z_3 = \langle \{ l_4, l_7 \}, \{ l_6, l_7 \}, \{ l_4 \}, l_6 \rangle, \]

where
\[ W_{7,6} = \text{“all distances are calculated”}, \]
\[ W_{7,7} = \neg W_{7,6}. \]

The token \( \alpha \) obtains characteristic

“the minimal distance with the corresponding class index”

in place \( l_6 \) and

“estimation of the current distance”

in place \( l_7 \).

Token \( \gamma \) enters place \( l_8 \) with initial (and final) characteristic

“\( \mu \text{threshold}, \nu \text{threshold}, \tau \text{threshold} \)”.

\[ Z_4 = \langle \{ l_6, l_8, l_{12} \}, \{ l_9, l_{10}, l_{11}, l_{12} \}, \{ l_6 \}, l_6 \rangle, \]

\[ l_9 \]

false true false true, \( \lor (l_4, l_7) \),

\[ l_{12} \]

\[ W_{12,9} \]

false \( W_{12,11} \)

\[ W_{12,12} \]

\( \land (\lor (l_6, l_{12}), l_8) \).
where
\[ W_{8,10} = W_{12,11} = -W_{12,12} \land (\mu_{\text{max}} > \mu_{\text{threshold}}) \land (\nu_{\text{resp}} < \nu_{\text{threshold}}), \]
\[ W_{12,9} = -W_{12,12} \land ((\mu_{\text{max}} > \mu_{\text{threshold}}) \land (\nu_{\text{resp}} \geq \nu_{\text{threshold}}) \lor (\mu_{\text{max}} \leq \mu_{\text{threshold}}) \land (\nu_{\text{resp}} < \nu_{\text{threshold}})), \]
\[ W_{12,12} = \text{“parameters } \mu \text{ and } \nu \text{ are not calculated for all classes”}, \]
where \( \nu_{\text{resp}} \) is the degree of nonmembership of the pattern to the class for which the degree of membership is equal to \( \mu_{\text{max}} \).

In place \( l_{12} \), where it will make one or more cycles, token \( \alpha \) obtains the characteristic

“evaluation of \( \mu_{\omega_i} \) and \( \nu_{\omega_i} \),”

where both numbers can be obtained in different ways. For example, we can use formulae

\[ \mu_{\omega_i}(x_u) = \exp(-d_i), \]
\[ \nu_{\omega_i}(x_u) = \exp(-d_i^{\text{next}}), \]

where \( i = 1, 2, \ldots, M \). When \( \nu_{\omega_i}(x_u) \) is evaluated for the nearest class, i.e., for the class corresponding to \( d_{\text{min}} \) (the first member of the sequence \( D \)), then \( d_i^{\text{next}} \) is the second distance in the sequence \( D \) and \( \vec{\pi}_i^{\text{next}} \) is the prototype vector of this, competitive class. In case, that \( \nu_{\omega_i}(x_u) \) is evaluated for all other classes, then \( d_i^{\text{next}} \) is the antecedent distance in the sequence \( D \), i.e., the distance to the competitive class. Other formulae can be applied instead of the above ones, as well. For example, the GN-model can use formulae

\[ \mu_{\omega_i}(x_u) = \frac{\| x_u - \vec{x}_i \|^{-\frac{2}{q}}}{\sum_{k=1}^{M} \| x_u - \vec{x}_k \|^{-\frac{2}{q}}}, \]
\[ \nu_{\omega_i}(x_u) = \frac{\| x_u - x_i^{\text{next}} \|^{-\frac{2}{q}}}{\sum_{k=1}^{M} \| x_u - \vec{x}_k \|^{-\frac{2}{q}}}, \]

where \( q > 1 \) is the degree of fuzziness (fuzzy generator).

Token \( \alpha \) obtains in place \( l_9 \) characteristic

“moving the pattern to each of the classes according to formula
\[ x_{u,j}^{\text{new}} = x_{u,j}^{\text{old}} (1 + a_{j,k} \tanh(l_{i,j})), \text{ for every } j = 1, 2, \ldots, n \text{ and } i = 1, 2, \ldots, M”, \]
where
\[ a_{j,k} = \frac{1}{1 + e^{-r_{j,k}}} = \text{sigmoid}(r_{j,k}), \]
\[ l_{i,j} = \frac{\bar{x}_{i,j} - x_{u,j}}{\sigma_{i,j}}, \]
where \( r_{j,k} \) is the correlation coefficient between the \( j \)-th coordinate (current feature under correction) and the \( k \)-th coordinate, which is the closest \( Xu \) feature to the class mean vector \( \bar{x}_i \); \( \sigma_{i,j} \) is the standard deviation of the \( j \)-th feature within the class \( \omega_i \); \( \bar{x}_{i,j} \) is the \( j \)-th coordinate of the class mean vector \( \bar{x}_i \),
\[ \tanh(l_{i,j}) = \frac{e^{l_{i,j}} - e^{-l_{i,j}}}{e^{l_{i,j}} + e^{-l_{i,j}}}. \]

Tokens \( \alpha \) and \( \gamma \) do not obtain any characteristics in places \( l_{11} \) and \( l_{10} \), respectively.

\[ Z_5 = \langle l_9, l_{13}, l_{14}, l_9, l_{13} \rangle, \]
where
\( W_{13,13} = -W_{13,14} \),
\( W_{13,14} = \text{“all distances are calculated”} \).

Token \( \alpha \) will cycle in place \( l_{13} \) and on the current (s-th) cycle it will obtain the characteristic

“estimation of the current distance using formulae
\[ d_i = \| x_u^s - \bar{x}_i \| = \sqrt{\sum_{j=1}^{M} (x_{u,j}^s - \bar{x}_{i,j})^2} \quad (i = 1, 2, ..., M), \]
where \( x_{u,j}^s \) is the moved pattern for the fixed \( s \) \( (s = 1, 2, ..., M) \).

Finally, it will enter place \( l_{14} \) with the characteristic

“the minimal distance with the corresponding class index \( s^{\min} \),
where \( s^{\min} \) is the index of element \( d_i \in D \) for which \( d_i = d_i^{\min} \).

\[ Z_6 = \langle l_{14}, l_{15}, l_{16}, l_{14} \rangle, \]
where
\( W_{14,15} = \text{“not all distances has been calculated”} \),
\( W_{14,16} = -W_{14,15} \land (\mu_{\max} > \mu_{\text{threshold}} \land (\nu_{\text{resp}} < \nu_{\text{threshold}}). \)

The token \( \alpha \) enters places \( l_{15} \) and \( l_{16} \) without a new characteristic.
4 Conclusion

All this means that the location correction actually changes $\mu_{\omega_i}$, $\nu_{\omega_i}$ and $\pi_{\omega_i}$ ($= 1 - \mu_{\omega_i} - \nu_{\omega_i}$) at each iteration and also means that operator $F_{\alpha,\beta}(x)$ (see Fig. 3) is indirectly applied, i.e. $X_u$ is moved within triangle $OAB$ (in fact $XCD$) towards $AB$ (actually $CD$) and parameters $\alpha$ and $\beta$ of operator $F$ are functions of $r$ and $l$, i.e.

$$\alpha = f_1(r_{j,k}, l_{j,k}),$$
$$\beta = f_2(r_{j,k}, l_{j,k}).$$

Obviously, the moving procedure stops when $\pi_{\omega_i}(x_u^{\text{new},i} < \pi_{\text{threshold}}(x)$, i.e. it is smaller than a preset threshold. In [6] it is proved that such an algorithm is convergent.

An intuitionistic fuzzy version of the nearest prototype classification method is proposed. That version is based on a procedure of moving of the pattern towards the classes and it combines the statistical and fuzzy approaches. Thus, the values of $\mu_{\omega_i}(x_u)$ and $\nu_{\omega_i}(x_u)$ have been adjusted and made more precise, and $\pi_{\omega_i}(x_u)$ has been considerably decreased. As a result the confidence in the classification decision is enhanced. During moving of the pattern, location can be changed in such a way that the initially nearest class may turn out to be a remote one. The procedure stops when the degree of indeterminacy is made negligibly small. The experimental results show that the combination of statistical and fuzzy approaches in pattern recognition leads to an improvement of the recognition accuracy, and it seems promising to implement such classification rules in some consulting systems for medical application where the confidence in the classification decision should be strong. The GN-model gives more light and understanding of this method and the possibility to be easily integrated in a complex decision support system.

![Fig. 3: An intuitionistic fuzzy operator $F_{\alpha,\beta}$](image)
References


