

Clasificación inteligente sinergia entre lógica difusa y algoritmos genéticos

José Adrián Rodríguez González

Octubre 2024

1. Introducción

Para el mes de octubre se leyó la literatura acerca de los algoritmos genéticos y la lógica difusa propuesto por (Ishibuchi et al., 1997), en este primer articulo menciona varios aspectos importantes de implementación de estos métodos. A su vez que en (Ishibuchi et al., 1995), explica otra metodología de aplicación para crear seleccionar reglas difusas.

2. Métodos de selección de sistemas difusos.

2.1. Método Michigan.

Primeramente, se crea un conjunto difuso A que contiene n subconjuntos difusos, que dependerán de x_i características de un conjunto de datos, siendo

$$A = \{A_i | A_i \subseteq A\} \wedge A_i = \{(x_i, \mu(x_i)) | x_i \in X\}$$

Los subconjuntos difusos se puede definir como (tabla 1):

Nivel	Valor
Very Low	1
Low	2
Medium	3
High	4
Very High	5

Figura 1: Tabla de conjuntos difusos enumerados

Después, se debe de tener un vector V , cada indice i del vector es una característica x del conjunto de datos, y los elementos del vector son un conjunto difuso A_i .

$$v_1 = [2, 1, 3, 5]$$

Esto significa, que los valores 2,1,3 y 5 son subconjuntos difusos de A , y la posición 0,1,2,3 son las características, por lo tanto, a la característica 0 se le evalúa con el segundo subconjunto difuso, la característica 1, es el primer subconjunto difuso, y de esa forma, se pueden definir reglas, tales como

if x_1 is low and x_2 is very low and x_4 is medium and x_5 is very high
then is Class 1 with CF:0,83

El CF es el grado de confianza de que sí es la clase dicha, para ello, se deben de realizar los siguientes cálculos:

Se calcula

$$\beta_{CT} = \sum_{x_{p1} \in CT} \mu_i^K(x_{pq}) * \mu_j^K(x_{p2})$$

, De forma generalizada, ya que esta manera de expresarlo es en base a dos características.

$$\beta_{CT} = \sum_{x_{p1} \in CT} \prod_{q=0}^Q \mu_i^K(x_{pq})$$

Siendo q una característica y Q el conjunto de características que $q \subseteq X_q$. Se multiplica cada una de las funciones de activación dadas por la regla de todas las características seleccionadas de cada instancia del conjunto de datos. CT es una clase específica, por lo que debemos de separar nuestro conjunto de datos por las clases que existan, de esa manera, la sumatoria de las multiplicaciones de los subconjuntos difusos de cada característica será el β_{CT} de esa clase. Si un $\beta_{CT} = 0$, eso significa que esa regla es incompatible para esa clase.

Al conjunto de β_{CT} , se busca aquel que tenga el máximo valor, de tal manera que:

$$\beta_{CX} = \max\{\beta_{CT1}, \beta_{CT2}, \dots, \beta_{CTM}\}$$

Si dos o más clases tienen el mismo valor o todos valen 0, entonces C_{ij}^k no puede ser determinado, por lo que este tipo de reglas se les considera *dummies*, se les suele llamar como ϕ , sino $C_{ij}^k = CX$. Ya que se pudo generar esa clase, entonces se calcula el grado de confianza de esa clase.

$$CF_{ij}^K = \frac{(\beta_{CX} - \beta)}{\sum_{T=1}^M \beta_{CT}}$$

Donde

$$\beta = \frac{\sum_{T=1, T \neq X}^M \beta_{CT}}{(M - 1)}$$

De esa forma, podemos encontrar la tasa de confianza que posee una regla dada a una clasificación.

Para poder clasificar un patrón, o instancia, se genera un S conjunto de reglas, siendo $R_{ij} \subseteq S$, después, se calcula el $\beta_{CT} * CF_{ij}^K$ de la instancia x_p , para todas las $R_{ij} \subseteq S$, de tal forma, que se encuentra el valor máximo de

este calculo de todo el conjunto de reglas y la clase asociada a la regla, será la clasificación del patrón.

$$\alpha_{CT} = \max\left\{\prod_{q=0}^Q \mu_i^K(x_{pq}) * CF_i^k | C_i^k = CT, R_i^k \in S\right\}$$

De esta forma $\alpha_X(CX)$ es

$$\alpha_{CX} = \max\{\alpha_{C1}, \alpha_{C2}, \alpha_{C3}, \dots, \alpha_{CM}\}$$

Selección

La selección para este modelo de acuerdo a lo propuesto por (Ishibuchi et al., 1997), se hace una ruleta de probabilidades.

$$P(R_j) = \frac{f(R_j) - f_{\min}(S)}{\sum_{R_i \in S} \{f(R_i) - f_{\min}(S)\}}$$

Cruza y mutación

Acorde a (Fig.2), la mutación es una cruce de los cromosomas de los individuos seleccionados, y la mutación es un ligero cambio que se realiza en el gen de un individuo. La pieza clave para Michigan, es que el algoritmo genético se

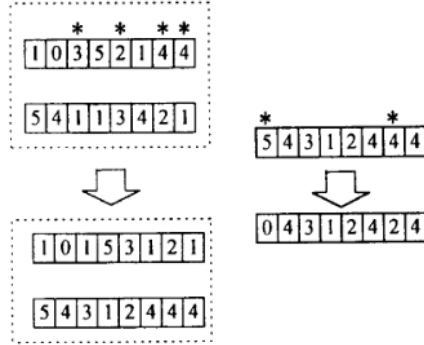


Figura 2: Cruza y mutación

emplea a cada regla del conjunto.

2.2. Método Pittsburgh.

Pittsburgh emplea el mismo proceso, solamente, que en vez de realizarlo en base a reglas individuales, lo hace por subconjuntos de reglas, de tal manera

que tanto las formulas de evaluación, como de selección, se hace en base a un conjunto de reglas

Para escoger el método, dependerá del tipo de problema que estemos manejando, no siempre se usan los mismo métodos para todos los problemas, ya que en algunas ocasiones, no resultan en un sistema óptimo.

3. Trabajo del mes

Como se puede observar, se recopiló la información de Ishibuchi, sin embargo, también se realizó un primer modelo en un problema de clasificación binaria como el conjunto de datos de cancer de mamá de Wisconsin.

4. Partes del problema.

Se leyó el conjunto de datos, y se probó con un modelo inicial de Random forest y se observó que para es conjunto de datos obtuvo un 95 % de accuracy. Sin embargo, para poder comenzar con los conjuntos difusos, se divide el conjunto de datos en un 70 % de entrenamiento y un 30 % de prueba, todos los datos fueron normalizados. Tanto la entrada del conjunto de entrenamiento como la salida, se dividió en 5 subconjuntos difusos.

Finalmente se realiza el algoritmo genético, en base a Michigan:

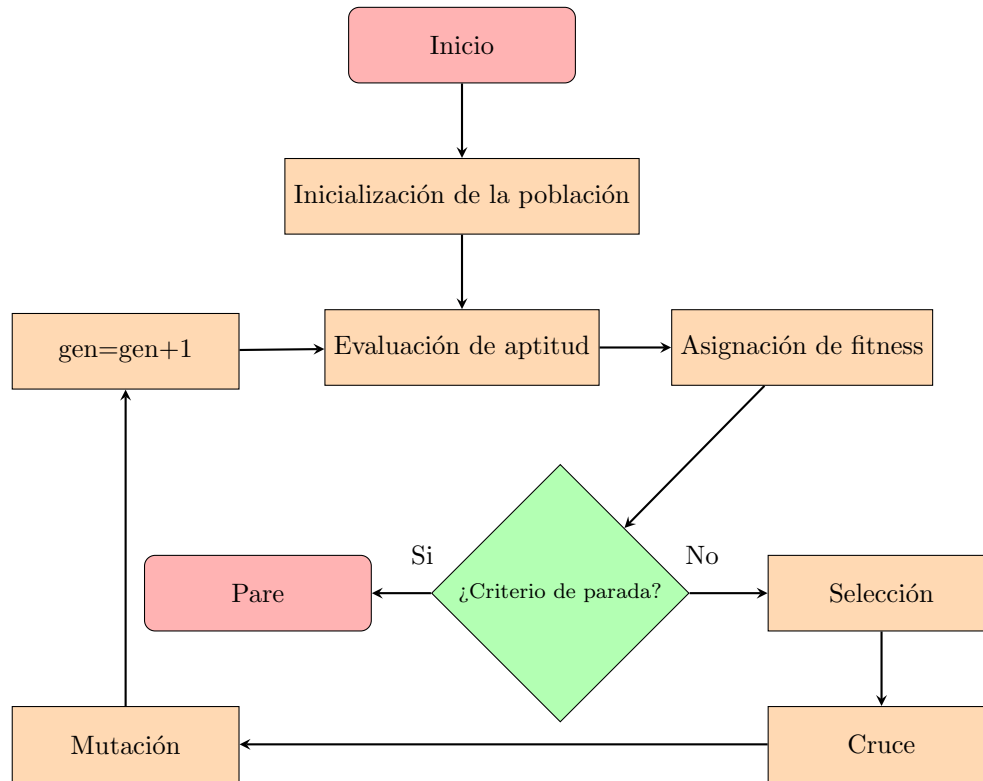


Figura 3: Diagrama de flujo

El proceso de selección, para este caso se utilizó una forma elitista. Sin embargo, una de las mayorías problemáticas que se presentaron para el primer mes fue que el clasificador, se ponderaba para una sola clase todas sus predicciones, por ejemplo, si en la primera clase, predijo que era 1, todas las predicciones tendían a 1, y si eran 0, todas tendían a 0, por lo que para el mes siguiente se realizaran las correcciones de la predicción.

Referencias

- Ishibuchi, H., Nakashima, T., and Murata, T. (1997). Comparison of the michigan and pittsburgh approaches to the design of fuzzy classification systems. *Electronics and Communications in Japan Part Iii-fundamental Electronic Science - ELECTRON COMMUN JPN III*, 80:10–19.
- Ishibuchi, H., Nozaki, K., Yamamoto, N., and Tanaka, H. (1995). Selecting fuzzy if-then rules for classification problems using genetic algorithms. *IEEE Transactions on Fuzzy Systems*, 3(3):260–270.