

# Aprendizaje Inteligente: Métodos de Mountain Clustering y Participatory Learning

**Carmen López Martínez**

**Programa de Doctorado: Diseño, Análisis y Aplicaciones de Sistemas Inteligentes**

**Curso: "New Trends on Intelligent Systems and Soft Computing"**

**klopez@ugr.es**

En este documento se presenta un resumen de la conferencia "Methods for Intelligent Learning" impartida por el Prof. Ronald R. Yager en el Seminario sobre nuevas tendencias en Sistemas Inteligentes y Soft Computing, celebrado en la E.T.S.I. Informática de la Universidad de Granada en octubre de 2003.

## 1. Introducción

Los métodos de aprendizaje que se presentan a continuación están basados en la forma de aprender que tenemos los humanos. Mountain Clustering está inspirado en como localizamos el centro de las agrupaciones de objetos similares de forma visual. Por su parte, el método Participatory Learning se basa en la forma en que vamos modificando nuestro sistema de creencias a lo largo de nuestra vida. Estos cambios se realizan poco a poco y poniendo nosotros de nuestra parte, rechazando en principio, cualquier conocimiento que choque de plano con lo que nosotros suponemos cierto.

## 2. Mountain Clustering

Mountain Clustering es un método usado para determinar el centro de los cluster para la clasificación de los elementos de un conjunto. El proceso se puede dividir en tres fases. En una primera fase, creamos una partición del espacio que determinará un subconjunto de puntos que constituirán los potenciales centros de los clusters. A continuación construiremos la función en montaña usando los datos de la muestra. Por último, seleccionaremos los centros de los cluster, escogiendo aquellos para los que exista un máximo de la función y eliminaremos el efecto de ese máximo en la función.

### 2.1. Discretización del espacio

Seleccionaremos un subconjunto finito de puntos del espacio de definición de los datos. Estos puntos serán los centros potenciales de los clusters. Este subconjunto se puede seleccionar de diferentes maneras: aleatoriamente, mediante puntos difusos, usando conocimiento previo sobre el subconjunto, creando una rejilla (que puede ser uniforme o no) sobre el espacio de definición de los datos, etc.

La aproximación de los centros seleccionados a los centros reales es muy sensible a la elección de los puntos iniciales. Sin embargo, cuanto más puntos seleccionemos, más tiempo de computación será necesario para la construcción de la función en montaña. Por otro lado, cuanto más fina sea la selección de los puntos, mejor será nuestra aproximación a la solución.

## **2.2. Construcción de la Función en Montaña**

Construiremos la función en montaña definida sobre el conjunto de los nodos de la rejilla añadiendo a cada nodo, para cada dato del conjunto de muestra, una cantidad proporcional a la distancia al nodo. Cuanto más cercano esté un dato al centro, mayor será su aportación al valor final de la función en el nodo. Cuanto mayor sea el valor de la función en montaña en el nodo, mayor será la probabilidad de ese nodo de ser elegido como centro de un cluster.

## **2.3. Selección de Centros de los Clusters**

El objetivo del método es localizar los centros de los clusters. Para ello localizaremos el máximo global de la función en montaña que hemos construido en el paso anterior. Si hay más de un máximo global, elegimos uno aleatoriamente. Para poder continuar eligiendo centros, tenemos que eliminar el efecto de ese máximo en la función. Restaremos a la función en montaña un valor en cada nodo que será proporcional a la distancia del punto desde el máximo y proporcional al valor maximal. Repetiremos la operación iterativamente hasta alcanzar la condición de parada. Ésta se puede definir de diferentes maneras: Cuando la función en montaña esté destruida, es decir, su valor está por debajo de un parámetro que determina que la función ha sido totalmente desmontada, cuando hayamos obtenido el número de clusters definido al inicio o cuando los clusters encontrados definen un recubrimiento del espacio.

## **2.4. Generación de Reglas Difusas con Mountain Clustering**

Generaremos un sistema basado en reglas difusas cuyas reglas están basadas en esquemas del tipo IF entrada es X entonces salida es Y. Mountain Clustering se puede usar para proporcionar estructura a la base de conocimiento y estimar los valores iniciales de las funciones de pertenencia de los antecedentes y los consecuentes de las reglas. Usaremos las entradas/salidas del sistema para calcular los centros de los clusters. Para cada cluster  $C_i=(a_i, b_i)$  generaremos una regla del tipo IF entrada es  $A_i$  entonces salida es  $B_i$ , donde  $A_i$  y  $B_i$  son conjuntos difusos centradas en  $a_i$  y  $b_i$ . Podemos usar diversos métodos para refinar nuestro sistema de reglas, por ejemplo, el método del Gradiente Descendente.

# **3. Participatory Learning**

Un concepto fundamental en el modelo Participatory Learning es que nuestras creencias actuales influyen determinantemente en lo que estamos dispuestos a aprender. En el modelo de Participatory Learning en un primer nivel, cualquier información que sea incompatible con nuestras creencias será rechazada, clasificandola como errónea. Participatory Learning asume que nuestras creencias son correctas y cree que todo lo que no sea compatible, es erróneo. Por otro lado, las observaciones

que son compatibles con nuestras creencias, aunque algo diferentes, se usan para realizar pequeñas modificaciones en nuestro sistema.

Intentamos aprender un conjunto de variables de interés  $V(i)$ . Nuestro conocimiento del entorno nos viene representado como un conjunto de vectores  $D(i)$  que representan las observaciones medidas. El mecanismo que usaremos para actualizar nuestras creencias es:

$$V(i+1) = V(i) + \alpha \rho_i (D(i) - V(i))$$

$V(i)$  representa nuestro sistema de creencias actual. En cada momento, si nos llegan datos compatibles, será modificado.

$\alpha$  es el porcentaje de aprendizaje primario.

$\rho$  constituye el grado de compatibilidad de la observación actual con respecto a nuestro sistema de creencias. Se define como:

$\rho_i = F(D(i), V(i))$ , donde  $F(x, y)$  es una medida de distancia. Si  $\rho_i$  es igual a 0, los datos y las creencias son incompatibles, por lo que el sistema no aprende. En el caso en que  $\rho_i$  sea igual a 1, los datos y las creencias son totalmente compatibles, pero en este caso tampoco se produce ningún aprendizaje, puesto que los datos y nuestro conocimiento son iguales. En consecuencia, se produce el mayor grado de aprendizaje cuando las observaciones y nuestras creencias son muy similares, pero no exactamente iguales.

Una posible mejora del sistema consiste en incluir un índice de estimulación. Este índice refleja el número de veces que las observaciones del entorno han sido incompatibles con nuestras creencias. De esta forma, si notamos que durante mucho tiempo nuestras creencias han estado en oposición con los datos de entrada, el sistema considerará cambiar su conocimiento actual. Cuanto más alto sea el índice de estimulación, menos confiaremos en nuestro actual sistema de creencias, y más predispuestos estamos a cambiar el sistema actual.

## **4. Conclusiones**

### **4.1. Mountain Clustering**

Este método es muy dependiente del conjunto inicial de nodos seleccionados. Si los puntos elegidos son buenos (cerca de los centros de los clusters) las soluciones obtenidas serán buenas, y malas en caso contrario. Esto es debido a que no se realiza exploración de soluciones. Para asegurarnos una buena selección debemos partir de un gran conjunto de nodos, lo que incrementará el tiempo de ejecución.

Además se nos presenta el problema de la selección de los valores de los parámetros que regulan el comportamiento del algoritmo. Podemos ajustarlos mediante la ayuda de un experto, pero puede ser

una tarea complicada encontrar los valores óptimos.

Una aplicación de este método puede ser la selección de un conjunto inicial de clusters para refinarlo mediante otro método de agrupación, por ejemplo el algoritmo de las k-medias o el ISODATA.

## 4.2. Participatory Learning

El problema principal que presenta el Participatory Learning es la inicialización de la base de conocimiento. Dependiendo de lo acertada que esta sea, el resultado final será más o menos adecuado debido a que el método es muy conservador. Las ideas que se contradicen con nuestras creencias actuales son rechazadas en primera instancia, por lo que los cambios en el conocimiento son mínimos. Es posible que observaciones que sean correctas sean rechazadas porque entran en conflicto con nuestro sistema de creencias actual y no se presentan con la suficiente insistencia para ser aceptadas por el método de aprendizaje.

En general consiste en un método de aprendizaje lento, más adecuado para refinar una base de conocimiento obtenida a partir de otro método, como pueden ser los métodos de aprendizaje evolutivos o los basados en refuerzo, que para generar una base de conocimiento desde el principio.

## 4.3. Preguntas

- Sería interesante conocer si estos métodos se han usado en aplicaciones reales y con qué resultados.
- ¿Qué porcentajes de clasificación se han conseguido con el algoritmo de Mountain Clustering?
- ¿Qué valores de los parámetros para Mountain Clustering y para Participatory Learning han resultado más adecuados? ¿Cómo se ha llegado a esos valores?

## Bibliografía

- [1] *Approximate Clustering Via the Mountain Method*, Ronald R. Yager, Dimitar P. Filev, *IEEE Transactions on Systems, Man and Cybernetics*, 24, 1279-1284, Agosto 1994.
- [2] *Generatio of Fuzzy Rules by Mountain Clustering*, Ronald R. Yager, Dimitar P. Filev, *Journal of Intelligent and Fuzzy Systems*, 2, 209-219, 1994.
- [3] *A Model of Participatory Learning*, Ronald R. Yager, *IEEE Transactions on Systems, Man and Cybernetics*, 20, 1229-1234, Septiembre/Octubre 1990.
- [4] *A Fuzzy Logic View of Participatory Learning*, Ronald R. Yager, D. P. Filev, *Proceedings of the 2nd IEEE International Conference on Fuzzy Systems*, 912-917, 1993.
- [5] *Autonomous Mountain-Clustering Method Applied to Fuzzy Systems Modeling*, *Laboratório de Mecatrónica. Instituto Superior Técnico (I.S.T.), Portugal*,: <http://citeseer.ist.psu.edu/351988.html>.