

APRENDIZAJE MEDIANTE ÁRBOLES DE DECISIÓN

María C. García-Alegre, Javier Rivas, David Martín y Matilde Santos

A10.1 INTRODUCCIÓN

A continuación se describe una aplicación donde el sistema de aprendizaje automático diseñado sirve de ayuda al operario en la toma de decisiones para la vigilancia y control de una presa en una central hidroeléctrica. En el sistema desarrollado se representa el conocimiento experto mediante un árbol de decisión basado en lógica borrosa. A partir del análisis de las variables medioambientales y del nivel del agua embalsada, obtenidas mediante sensores *ad-hoc*, el sistema tiene por objetivo detectar situaciones precríticas, generar alertas tempranas, y determinar la apertura de las compuertas de desagüe.

En un embalse se llega a una situación de precrisis cuando el nivel del agua se aproxima al nivel superior del dique de contención del embalse, pues existe riesgo de que se produzcan vertidos por coronación, con posibles daños tanto en el muro como en los cultivos y la población civil aguas abajo. Los datos analizados, 5069 casos con muestreo horario, han sido proporcionados por Unión Fenosa

(Madrid) dentro del proyecto CENIT_TIC_2005 del CDTI: Homeland Security. Corresponden a un período en el que se produjeron varias avenidas en una presa de una central hidroeléctrica situada en el cauce del río Miño.

La aplicación desarrollada en MATLAB (<http://www.mathworks.com/>) integra el método de aprendizaje mediante un árbol de decisión fuzzy y visualiza los resultados del razonamiento para distintos casos. El objetivo de este prototipo es mostrar el funcionamiento del algoritmo basado en árboles de decisión fuzzy, ante problemas concretos de toma de decisiones utilizando conocimiento experto.

En esta versión se ha reducido la dimensión del problema, mostrada en el capítulo 10, a las tres variables que más influyen en la decisión a las que se ha asignando diferente valor de prioridad.

Los atributos analizados, en orden decreciente de prioridad son:

- *Lluvia*: Volumen de agua recogida (l/m^2)
- *Aforo*: Volumen de agua procedente del río que llega al embalse (m^3/s).
- *Cota*: Altura del agua en la presa respecto a la base del dique (m).

En este ejemplo, los conjuntos borrosos (fuzzy) o términos lingüísticos utilizados para resumir el universo de discurso de la variable *Lluvia* son {No, Media, Alta}. Estos conjuntos se representan mediante funciones de pertenencia trapezoidales en la Figura A10.1.

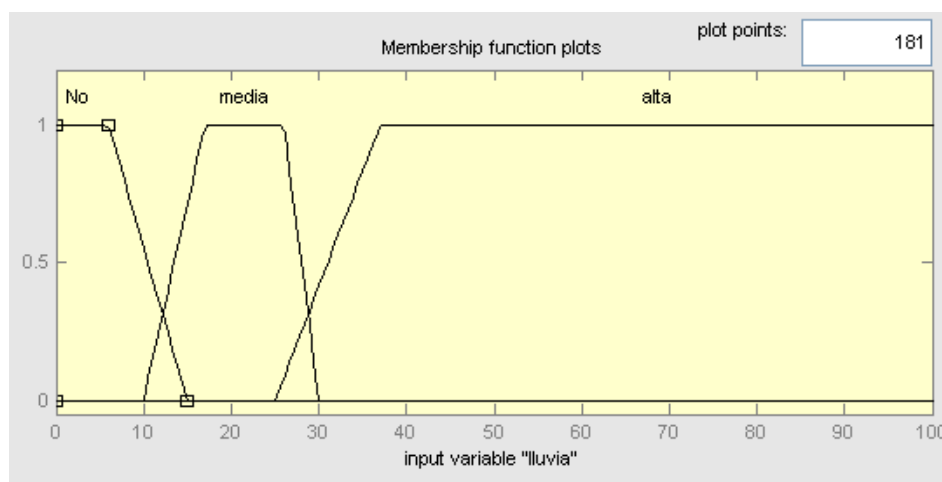


Figura A10.1. Funciones de pertenencia trapezoidales asociadas a los conjuntos borrosos de la variable Lluvia

De igual forma, los valores asignados a la variable *Cota* son {Poco, Medio, Mucho} y a la variable *Aforo* {Bajo, Medio, Alto}, Figura A10.2 y Figura A10.3.

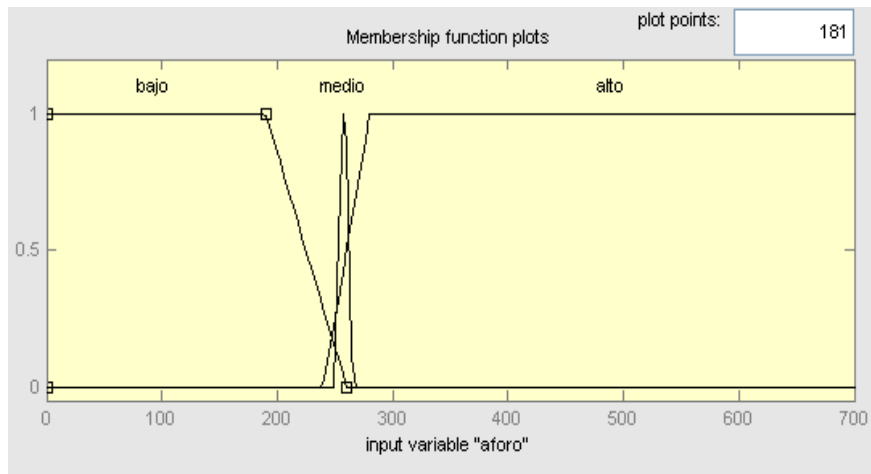


Figura A10.2 Funciones de pertenencia asociadas a los conjuntos borrosos de la variable Aforo

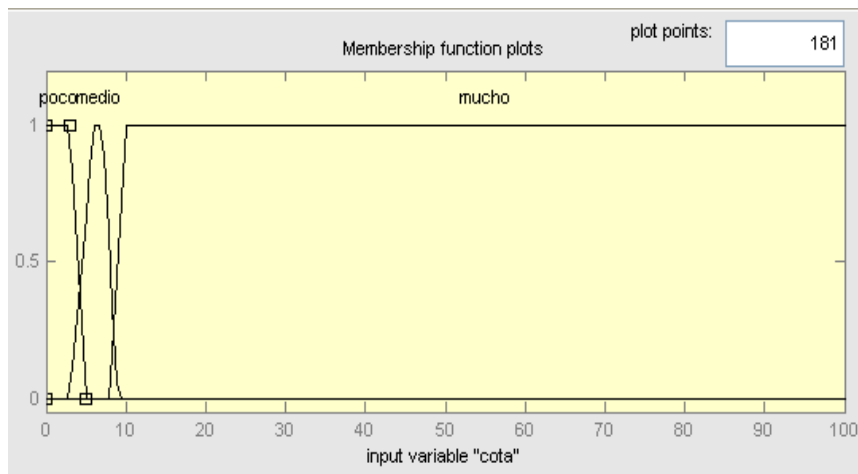


Figura A10.3 Funciones de pertenencia asociadas a los conjuntos borrosos de la variable Cota

En la Figura A10.1, Figura A10.2 y Figura A10.3, se observa el intervalo de valores numéricos que pueden tomar las variables, y los límites de los conjuntos

borrosos asociados a cada una de ellas: $Lluvia = \{\text{No}, \text{Media}, \text{Mucha}\}$, $Aforo = \{\text{Bajo}, \text{Medio}, \text{Alto}\}$ y $Cota = \{\text{Baja}, \text{Media}, \text{Alta}\}$.

La aplicación que se presenta a continuación está desarrollada en MATLAB y consta de un programa de tipo script principal denominado *Fuzzy_Presa*. Tras su ejecución se muestra el interfaz de usuario de la Figura A10.4.

A10.2 APRENDIZAJE

En este apartado se describe el funcionamiento del ejemplo propuesto para la comprensión del modelo de razonamiento basado en árboles de decisión fuzzy. El prototipo que se ha desarrollado, recibe como entradas, a través del interfaz de usuario, los valores numéricos asignados a las tres variables. Para obtener resultados razonables y significativos, éstos valores deben encontrarse en el intervalo de posibles valores de cada variable. Los intervalos de valores admisibles, aparecen reflejados en los diagramas de barras del interfaz, donde también se muestran las unidades de cada variable.

Las salidas que se obtienen del sistema de ayuda a la toma de decisiones determinan el estado de riesgo de la presa y la actuación a realizar sobre las compuertas. Las diferentes salidas del sistema o clases se muestran en la Tabla A10.1.

Clase	Nivel de Riesgo	Acción sobre compuertas
1	No hay riesgo	Compuertas cerradas
2	Riesgo Bajo	Compuertas cerradas
3	Riesgo Medio	Compuertas cerradas
4	Riesgo Medio	Compuertas abiertas
5	Riesgo Alto	Compuertas abiertas

Tabla A10.1. Cinco clases de salida del sistema de aprendizaje

Al ejecutar por primera vez el “*script*” principal aparece un conjunto de ventanas, Figura A10.4, con los valores por defecto para las tres variables, indicados en la Tabla A10.2.

Lluvia (l/m^2)	Aforo (m^3/s)	Cota (m)
0.0	0.0	0.0

Tabla A10.2 Valores por defecto o de inicialización de la aplicación

Interfaz Reglas Presa

SISTEMA DE RAZONAMIENTO AUTOMÁTICO

VARIABLES DE ENTRADA

Lluvia (l/m2)

0 100 80 60 40 20 0

Aforo (m3/s)

0 600 400 200 0

Cota (m)

0 100 80 60 40 20 0

Reglas activadas del árbol de decisión

Decisiones

INICIAR ÁRBOL DE DECISIÓN

SALIR DE LA APLICACIÓN

Figura A10.4 Ventanas del interfaz de usuario y valores por defecto de las 3 variables

Para ejecutar la aplicación y obtener un resultado del proceso de razonamiento implementado, se presiona el botón: **INICIAR ÁRBOL DE DECISIÓN**. En ese instante, el sistema inicia el proceso de razonamiento, aplicando el algoritmo de toma de decisiones basado en un árbol de decisión fuzzy.

Por último, el sistema muestra al usuario a través del interfaz, por una parte, las reglas que se han activado desde la raíz al nodo terminal, y por otra, la

decisión propuesta, que implica la actuación pertinente sobre las compuertas de la presa, así como una estimación del nivel de riesgo. El botón situado en la zona inferior derecha del interfaz, permite salir de la aplicación.

A10.3 CLASIFICACIÓN

En este apartado se detalla el proceso de razonamiento que sigue el sistema para tomar una decisión relativa a una de las clases definidas en la Tabla A10.1, para tres casos diferentes. En el primer caso, los valores numéricos proporcionados por los sensores para las tres variables de entrada son:

Lluvia (l/m ²)	Aforo (m ³ /s)	Cota (m)
3.5	370.6	1.1

La variable o atributo de mayor relevancia o prioridad es la *Lluvia*; por ello es la primera que se evalúa y se corresponde con la raíz del árbol. Se calcula la pertenencia del valor de la *Lluvia* a los tres conjuntos borrosos {No, Media, Baja}, representados mediante funciones de pertenencia trapezoidal. Asociando al valor numérico medido la etiqueta lingüística para la cual la pertenencia es máxima, resulta un valor “No” para el atributo *Lluvia*. Por tanto, se expande el nodo *Lluvia* y se selecciona la rama correspondiente a *Lluvia* es “No”, Figura A10.5.

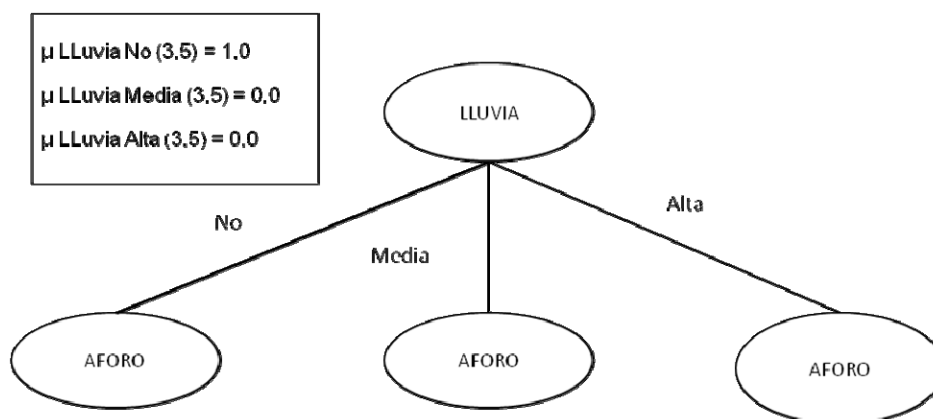


Figura A10.5 Expansión del Nodo Lluvia en el árbol de decisión

A continuación se evalúa la pertenencia del valor numérico medido del atributo *Aforo*, por ser la siguiente variable en prioridad, a los conjuntos borrosos {Poco, Medio, Mucho}. Se obtiene como resultado el valor “Mucho”. Por ello se expande el nodo *Aforo* y se selecciona la rama correspondiente a este valor, Figura A10.6.

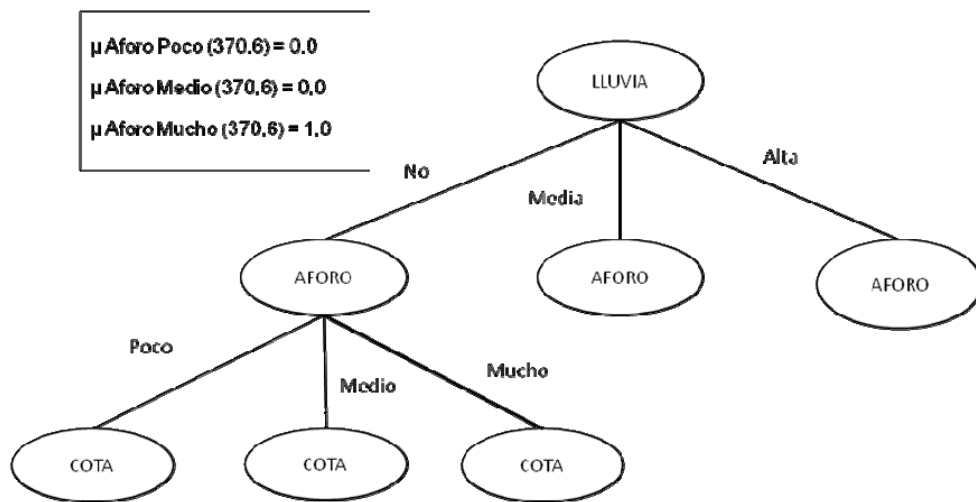


Figura A10.6 Expansión del Nodo Aforo en el árbol de decisión

En este caso se finaliza el proceso con la evaluación del grado de pertenencia del valor numérico de la variable *Cota*, 1.1 m, a los conjuntos borrosos {Baja, Media, Alta}, siendo el término “Baja” el resultado de la evaluación. De igual forma que en los casos anteriores, se expande el nodo *Cota*. Pero aquí, al haber llegado a los nodos terminales del árbol, se obtienen los resultados de la decisión, Figura A10.7.

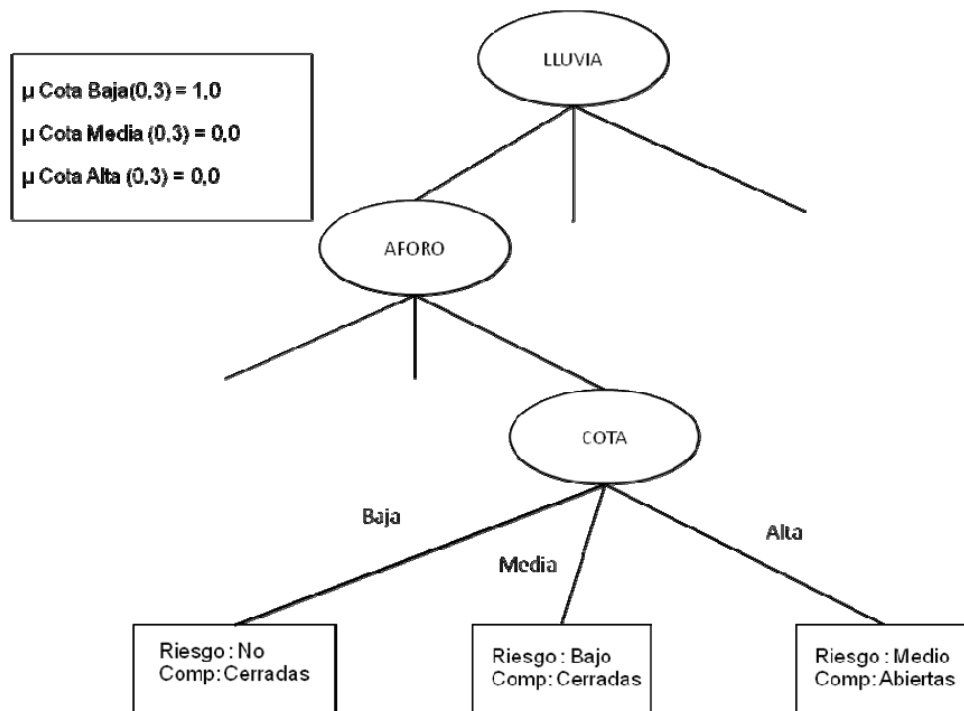


Figura A10.7 Expansión del Nodo Cota en el árbol de decisión

Con este ejemplo se muestra el proceso de inferencia seguido para llevar a cabo la toma de decisión o clasificación de un objeto o caso. Los nodos terminales contienen la estimación del nivel de riesgo y la actuación que se aconseja realizar sobre las compuertas de la presa. Finalmente, hacer notar que es la conjunción de todos los valores de las variables o atributos del antecedente de las reglas la que conduce al consecuente de la decisión o clase.

SI (*Lluvia* es “No”) AND (*Aforo* es “Mucho”) AND (*Cota* es “Baja”) ENTONCES (*Riesgo* es “No” AND *Compuertas* es “Cerradas”), Figura A10.8.



Figura A10.8 Interfaz de usuario mostrando los resultados del razonamiento para un caso

A continuación se muestran y comentan otros dos ejemplos, con el fin de ilustrar otras clases de salida del sistema. En este nuevo caso, los resultados se muestran en la Figura A10.9, y corresponden a los siguientes valores de las tres variables de entrada:

Lluvia (l/m ²)	Aforo (m ³ /s)	Cota (m)
20.0	370.6	8.0



Figura A10.9 Interfaz de usuario mostrando los resultados para el segundo caso

En este caso, la *Lluvia* es “Media”, *Aforo* es “mucho” y *Cota* es “Media”. Respecto al primer caso, se observa que aumenta la Lluvia y la Cota. Por ello, la activación de las Reglas: R2, R9, R26, conduce a otros grupos de nodos y la decisión sobre el nivel de riesgo es diferente. Se alerta de un riesgo catalogado como “Nivel Bajo” y no hay que abrir las compuertas.

Para finalizar, se presentan los resultados del tercer caso en la Figura A10.10, donde los valores de las variables de entrada son:

Lluvia (l/m^2)	Aforo (m^3/s)	Cota (m)
60.0	370.6	80.5

Interfaz Reglas Presa

SISTEMA DE RAZONAMIENTO AUTOMÁTICO

VARIABLES DE ENTRADA

Lluvia (l/m^2)
60.0

Aforo (m^3/s)
370.6

Cota (m)
80.5

Reglas activadas del árbol de decisión

- R3: SI Lluvia es Alta ENTONCES activa grupoLluviaAlta
- R12: SI Aforo es mucho ENTONCES activa grupoLluviaAltaAforoMu
- R36: SI Cota es Alta ENTONCES Estado es RiesgoAlto y Acción es

Decisiones

- Alerta Riesgo: nivel alto
- Abrir compuertas

INICIAR ÁRBOL DE DECISIÓN

SALIR DE LA APLICACIÓN

Figura A10.10 Interfaz de usuario, mostrando los resultados del proceso de razonamiento para el tercer caso

En este caso, la lluvia ha aumentado considerablemente, *Lluvia* es “Alta” y esto, como en el caso anterior, se ve reflejado en un fuerte aumento de la Cota, *Cota* es “Alta”, para un mismo valor del Aforo, *Aforo* es “mucho”. En este caso, la presa se encuentra en un estado de riesgo alto y lo más conveniente es abrir las compuertas para prevenir una crisis ante un posible riesgo de avenida, que podría dar lugar a vertidos por coronación, esto es vertidos por encima del dique de la presa.