

CAPITULO 5

Método propuesto:

Razonamiento Evidencial Dinámico

5.1. Introducción

Esta tesis presenta y describe un nuevo método de clasificación (RED: razonamiento evidencial dinámico) basado en el razonamiento evidencial al que se le introducen una serie de modificaciones.

RED permite la inclusión de nueva evidencia para el proceso de clasificación y propone una regla de decisión para la asignación de un objeto desconocido a una clase.

El algoritmo de razonamiento evidencial provee una forma de combinar evidencia a partir de distintas fuentes de información. Es una técnica de clasificación supervisada, que utiliza un conjunto de muestras de entrenamiento. RED agrega una etapa de aprendizaje para introducir la evidencia, lo que provoca una mejora en la precisión. Este proceso consiste en observar el valor obtenido de Khat cuando se realiza la clasificación del conjunto de muestras de entrenamiento, y decidir si es necesario modificar la evidencia actual.

La regla de decisión planteada utiliza, en principio, la medida de plausibilidad y la de soporte y estudia una serie de situaciones antes de tomar la decisión sobre la clase a la que se asignará el objeto desconocido. Su objetivo es realizar un análisis de la aceptación o no aceptación del objeto desconocido como perteneciente a una de las clases definidas para el algoritmo. Cuando no existe claridad respecto de esta decisión usando las medidas de soporte y plausibilidad, se agregan otras medidas como la cantidad de fuentes que aportan evidencia a favor de una clase y un peso para mostrar la “importancia” de la información de esa fuente para la clase en

cuestión. Con esto se quiere lograr obtener decisiones más cercanas a la realidad, analizando cuidadosamente los casos en que el soporte para varias clases es similar o cuando la evidencia no sea “suficiente”. Se consideran distintas alternativas para el significado de “suficiente” y diferentes posibilidades de decisión en dichos casos.

Se aplica el método propuesto para clasificar cultivos en imágenes hiperespectrales de la región de Nebraska (USA) [Pra98]. Se presentan los resultados obtenidos para evaluar la precisión de RED.

5.2. Descripción del método

La técnica de Razonamiento Evidencial propuesta por Peddle [Ped95 b], utiliza un conjunto de muestras de entrenamiento, del que se extraen las medidas de soporte y plausibilidad involucradas en la clasificación. Las fuentes de información a partir de las que se calculan las medidas de soporte y plausibilidad pueden ser de diferente naturaleza, por ejemplo se puede tener el resultado de una clasificación previa (rótulo asignado en la clasificación), un indicativo de la textura de una región, el valor de intensidad de un pixel, etc.

El método RED propone una forma de incorporar evidencia como parte de un proceso de aprendizaje en el que se va mejorando la precisión del clasificador. Con este fin, dado un conjunto de muestras, se evalúan los resultados de su clasificación y en caso de no ser "óptimos" se agregan los datos de las muestras (aquellos generados por las diferentes fuentes de información) como nueva evidencia. La forma en que se analizan los resultados de la clasificación para determinar su precisión, es a través de la matriz de error o confusión y el indicador Khat (ver **Capítulo 4, Sección 6** [Ros86]). Este indicador permite analizar el error ocurrido en la clasificación. Para poder determinar a qué llamamos “óptimo” se utiliza un parámetro de precisión α que pone un umbral para el aprendizaje.

En este sentido, RED puede ser considerado como una combinación del clasificador de razonamiento evidencial con un reconocedor de aprendizaje supervisado.

En una primera etapa, trabaja igual que el razonamiento evidencial. Se almacena información sobre las diferentes clases involucradas a partir de un conjunto de muestras iniciales, de las cuales se conoce a priori a qué clase pertenecen. A partir de esta información se calcula el soporte para cada muestra de una clase y fuente en particular (se sugiere tener presente lo explicado en el **Capítulo 4**).

La segunda etapa consiste en permitir que el clasificador evidencial aprenda de los resultados de la clasificación de diferentes conjuntos de entrenamiento, y se necesita ayuda del experto, de modo de incorporar nuevos datos que mejoran el comportamiento del mismo.

Esto está basado, como se dijo, en la idea de un clasificador con aprendizaje supervisado. Cuando se habla de reconocimiento con aprendizaje supervisado, se quiere indicar que existe un proceso *maestro* que se encarga de enseñar al sistema, corrigiéndole cada vez que se equivoque, es decir, cuando clasifique incorrectamente un objeto [Mar94]. Es necesario entonces, que este proceso maestro tenga una forma de evaluar cuando se está incurriendo en un error o cuando se ha tomado una decisión correcta.

Se tiene por ejemplo, un vector $x(\mathbf{k})$ (\mathbf{k} se refiere al instante genérico en que se extrae la muestra) del conjunto de muestras de entrenamiento, cuya pertenencia a una clase específica es conocida por el maestro.

La muestra $x(\mathbf{k})$ es clasificada por el reconocedor automático en virtud del conjunto de funciones discriminantes (ffdd) que en ese instante (\mathbf{k}) se esté manejando. Se modifican las funciones discriminantes (ffdd) en función de los resultados de la clasificación que se realizó.

La modificación se basa en si se produjo un acierto (coincide la decisión con la del maestro) o un error (difiere la decisión con la del maestro). Se pueden plantear diferentes algoritmos de modificación de las ffdd. Se puede utilizar para éstos un parámetro de error que dice si hay acierto o error, en nuestro caso se utiliza el parámetro de precisión α para determinar si se debe modificar la evidencia.

De esta forma, RED utiliza un conjunto de datos a clasificar, como muestras de aprendizaje, de las cuales se conoce a priori a qué clase pertenecen. De acuerdo a esta información y a los resultados obtenidos en la clasificación, se permite al clasificador modificar el conocimiento que tenía hasta el momento realizando un ajuste mediante la incorporación de nueva evidencia.

Cuando se decide incorporar nuevas muestras a la base de conocimiento del clasificador se puede estar dentro de dos posibilidades:

Dado el valor de una muestra de una cierta fuente y clase (se sugiere tener en cuenta las definiciones presentadas en el **Capítulo 4 – Sección 5**):

1. El valor puede ya existir teniendo un soporte mayor que cero.

En este caso se procede de la siguiente manera:

- Se incrementa la frecuencia de ocurrencia (**fo**) de la muestra
- Se actualiza el valor de cantidad de muestras (TSn) para la clase en cuestión

2. El valor puede no existir, no tiene soporte:

Entonces:

- Se agrega el valor de la muestra para la clase y fuente involucradas
- Se actualiza el TSn de la clase

Tanto para el caso 1 como para el 2, se debe cumplir con un tercer paso luego de haber procesado todas las fuentes involucradas:

- Recalcular el soporte para todos los valores de muestra de cada fuente dentro de la clase en cuestión.

Otra modificación introducida al método de razonamiento evidencial es una regla de decisión diferente, que toma en cuenta cuál es la clase con mayor soporte y analiza su aceptación o no como clase definitiva de acuerdo a los resultados de dos etapas de decisión precedentes: **a)** rechazo por falta de evidencia y **b)** rechazo por ambigüedad.

a) Rechazo por falta de evidencia

En esta etapa de decisión se evalúa la situación en la que la clase final a pesar de ser la de más soporte puede tener muy baja evidencia asignada, es decir la evidencia puede no ser “suficiente” para rotular al objeto como perteneciente a la clase. A esta decisión la llamaremos **rechazo por falta de evidencia**. Pero, inmediatamente surge la pregunta de cómo saber cuándo se tiene evidencia “suficiente” o a qué llamamos evidencia “suficiente”.

Regla de decisión con umbral para la evidencia

Se comenzó por introducir un valor de umbral para la evidencia, luego si la clase con mayor soporte estaba por debajo del umbral, se asignaba el objeto a la clase desconocida. Como suele ocurrir en estas situaciones la elección del umbral se basaba en relaciones empíricas encontradas en los datos. Los resultados obtenidos eran fuertemente dependientes de dicho umbral, invadiendo de muestras a la clase desconocida para casos de valores de umbral altos, o asignando, en algunas situaciones, incorrectamente el objeto a una clase para valores de umbral pequeños. Sin embargo, para algunos valores del umbral los resultados fueron razonables. Por ejemplo, entre las opciones elegidas se utilizó el máximo soporte encontrado entre todas las muestras (**Ec 5.1**), con lo cual se pedía tener un mínimo de evidencia que sea resultante de la combinación de varias fuentes para poder superar dicho umbral. De todas formas, se trató de buscar una alternativa en la que no se tuviera que utilizar un umbral para evitar la búsqueda del valor adecuado para este. En la siguiente subsección se detalla una posible alternativa.

(Ec 5.1) Sea S_j el soporte de la clase j que resultó ser el máximo encontrado entre el de todas las clases.

$S_j < T$ Clase final = desconocida

$S_j \geq T$ Clase final = Clase cuyo soporte es S_j

$T = \text{Max}(S_{ij})$ donde S_{ij} es el soporte de una muestra M_{ij} con $1 \leq i \leq N$ y $1 \leq j \leq M$

N : cantidad de fuentes y M : cantidad de clases

Regla de decisión considerando la cantidad de fuentes que aportan evidencia

Otra posibilidad analizada fue la de considerar qué cantidad de fuentes aportan evidencia en favor de la clase con mayor soporte. Qué sucede si en el ejemplo planteado en el **Capítulo 4, Sección 5**, sólo uno de los médicos aporta evidencia para la clase seleccionada por ser la de mayor soporte. ¿ Puede decirse que hemos optado por la decisión correcta?. Habría que estudiar la reputación del médico, cuán confiable es, en qué se basa su creencia, si es suficiente, etc. Pero, si fueran tal vez dos los que aportan evidencia tendríamos más seguridad en arriesgar por una decisión. Generalmente, es lo que una persona realiza en caso de un diagnóstico médico cuando se trata de casos complejos o dudosos, realiza una consulta con otro profesional para tener mayor seguridad. Y si ambos opinan lo mismo mayormente optamos por confiar en la opinión de éstos.

La regla de decisión que se probó entonces, analiza **cuántas son las fuentes que participan aportando evidencia para la clase con mayor soporte**. Para esta tesis se experimentó poniendo una restricción para que al menos dos fuentes hayan aportado evidencia para la clase con máximo soporte, en caso contrario se considera que el objeto debe ser asignado a la clase “desconocida”. No se tiene la evidencia “suficiente”. Además, a esta restricción se le agrega el caso de tener como mayor soporte un valor de 0, lo cual implica que no hay evidencia para la clase en cuestión, ni para ninguna de las restantes. También en tal situación la clase final seleccionada será la desconocida (**Ec 5.2**). Cabe aclarar que el valor de restricción para la cantidad de fuentes puede variarse de acuerdo a la aplicación particular con que se esté trabajando, exigiendo una mayor o una menor cantidad de fuentes que aporten evidencia a la clase de mayor soporte.

El cálculo de las medidas de soporte para cada clase se realiza mediante la combinación de los valores de soporte (Si) dados por cada fuente (i) a cada una de las clases (ver **Capítulo 4**).

(Ec 5.2) $C_j < R_f$ o $S_j = 0$ Clase final = desconocida

C_j = Cantidad de fuentes que aportan evidencia para la clase j (Clase con máximo soporte)

R_f = Valor de Restricción de Fuentes

b) Análisis de Ambigüedad

La segunda situación que se investiga es la que se produce si varias clases presentan un soporte similar al máximo. Esto implica que no hay una verdadera claridad respecto de la clase por la que se debe decidir. Para evaluar esto, se utiliza la medida de plausibilidad.

Se sabe que el intervalo de valores para la plausibilidad está entre 0 y 1 y dado que la plausibilidad indica cuanta evidencia hay sin asignar, y que podría llegar a ser asignada a la clase en cuestión, esto da una idea de cuanta evidencia se tiene distribuida entre las clases restantes. Si hay menos de un 50% de evidencia sin asignar, significa que hay un alto porcentaje que fue distribuido entre el resto de las clases. Esto implica que no hay una claridad en la decisión por la clase del objeto desconocido. Cuando se tiene el mayor valor de soporte igual a 1, entonces la plausibilidad deberá ser 0, y no hay evidencia asignada a las otras clases del marco de discernimiento. El caso opuesto sería tener como mayor soporte entre todas las clases un valor de 0, y entonces la plausibilidad sería de 1. De todas formas en tal situación, todas las clases tendrían igual chance y por esto se asigna el objeto a la clase desconocida. Cuando exigimos un 50% o más de plausibilidad, pedimos que al menos exista un 50% de evidencia que esté sin asignarse a las clases restantes, con lo cual la clase en cuestión (la de mayor soporte) tendría más chances de ser elegida como la definitiva.

Se estudiaron diferentes valores para el umbral. Queda de manifiesto a través de los estudios realizados sobre los valores de umbral para el porcentaje, que a mayor valor de umbral se impone una mayor restricción para arriesgar por una clase.

La otra idea estudiada dentro de la etapa de rechazo por ambigüedad, es la de evaluar la distancia existente entre los valores de evidencia de la clase con mayor soporte y el resto de las clases. Si alguna de las clases posee una distancia inferior o igual al parámetro de distancia se considera que no existe claridad para decidir entre las clases basándonos en el mayor soporte. El valor 0.01 parece ser razonable para la aplicación en cuestión y permite obtener los resultados esperados. Sin embargo, queda como trabajo futuro realizar un estudio más fundamentado para este parámetro.

Por otra parte, se plantea la pregunta de cuál es la decisión más acertada en caso de ocurrir la situación de tener que rechazar por ambigüedad (es decir es cierta alguna de las restricciones anteriores).

Una solución dada aquí fue asignar el objeto a la clase desconocida, para cuando se tiene un valor de porcentaje inferior al 50% o la distancia entre las clases está por debajo del umbral propuesto. En este caso, se ve que el algoritmo disminuye su precisión inclinándose mayormente por asignar los objetos desconocidos a la clase desconocida. Por lo que se prosiguió a buscar otra regla que permita decidirse por una de las clases propuestas.

La nueva opción adoptada para los casos en que se está por debajo del umbral es continuar analizando la elección y tratar de asignar el objeto a una clase entre las posibles (y no a la desconocida), teniendo en cuenta la cantidad de fuentes que aportan evidencia, la importancia de dichas fuentes para cada clase, y el valor de soporte obtenido (**Ec 5.3**). La combinación de estos tres factores permite tomar una decisión en situaciones de confusión. Para la cantidad de fuentes se considera un índice (**ICF**: Índice de Cantidad de Fuentes) que está dado por el valor de cantidad de fuentes sobre el total de las fuentes (esto da un número entre 0 y 1). Además se tiene un “peso” que se le asigna a cada relación fuente – clase (valor entre 0 y 1) que indica la importancia de la información de la fuente para la clase dada. Esto implica un conocimiento previo de la influencia de las fuentes para las diferentes clases. Para esta aplicación, por ejemplo se sabe que cuando se trata de distinguir entre lo que es

vegetación sana de lo que no lo es, las bandas del cercano infrarrojo y el rojo resultan de mayor interés al momento de decisión (ver **Capítulo 2, Sección 7**). Por último, el mayor valor obtenido de la combinación de estos tres aspectos permite determinar la decisión a tomar. La combinación está dada por la suma del índice de cantidad de fuentes y el soporte, multiplicada por el peso otorgado como lo indica la ecuación **Ec 5.3**.

Ec 5.3 Se selecciona la clase j correspondiente a $Max(\text{peso_total}_j * (ICF_j + S_j))$ con $1 \leq j \leq N$ donde N es el total de clases.

ICF : Índice de Cantidad de Fuentes para la clase j

S_j : Soporte para la clase j

peso_total_j : $\sum_{i=1..N} \text{peso}_{ij} / C_j$, donde i representa las fuentes que aportan evidencia para la clase j

$1 \leq j \leq M$ y C representa la cantidad de fuentes aportando evidencia para la clase j .

Regla utilizada por RED

Como consecuencia de los estudios realizados y explicados anteriormente, se determinó para RED la siguiente regla:

- a) **Análisis por falta de evidencia:** se evalúa la cantidad de fuentes que aportan evidencia a favor de la clase con mayor soporte (**Ec 5.2**). Para este estudio se puso como restricción tener al menos dos fuente, por lo que la regla asigna el objeto a la clase desconocida en caso de no cumplirse la restricción. De igual manera se procede si el mayor soporte encontrado es 0. Para este objetivo el algoritmo debe mantener la cuenta de cuántas fuentes aportan evidencia para cada clase (soporte mayor que cero).
- b) **Análisis de rechazo por ambigüedad:** si todavía no se optó por la clase desconocida se evalúa el rechazo por ambigüedad utilizando como umbral el 50% del intervalo posible de la plausibilidad (se podría decir que esto da un intervalo de confianza para la decisión). Cuánto más porcentaje se pida mayor será la evidencia que puede ser asignada y que no lo fue por la incertidumbre

asociada a los datos. En caso de no superarse dicho umbral o cuando existe alguna distancia entre la clase de mayor soporte y el resto de las clases menor que el parámetro de distancia, se estudia entre las clases posibles la que da mayor combinación entre soporte, peso de la fuente para la clase e índice de cantidad de fuentes, y se asigna el objeto a dicha clase (**Ec 5.3**).

Finalmente, cuando se clasifica una región homogénea, por ejemplo, tomar una porción de un campo cultivado cualquiera de la imagen, se introduce una última decisión o una etapa de post-clasificación, que se basa en la cantidad de píxeles asignados a cada una de las clases, si se encuentra que más del 80% de los píxeles fueron asignados a la misma clase, la región es considerada o rotulada como perteneciente a dicha categoría. En caso contrario, se le otorga el rótulo de “Desconocida”. Esta última fase del clasificador RED sólo se utiliza para el caso que se quiera dar al usuario final un valor único de clase para la región seleccionada, y esto agrega información. Aquellas regiones donde se desconoce la pertenencia a una clase, terminarán por asignarse a la clase desconocida.

5.3. Objetivos

El objetivo de la modificación introducida al razonamiento evidencial es, por un lado, agregar un proceso de aprendizaje para el clasificador, de manera tal que el mismo incorpore nuevas piezas de evidencia para las distintas clases, hasta lograr un comportamiento "óptimo" del mismo para el conjunto de entrenamiento con el que se trabaja.

La definición de "óptimo" depende de la aplicación en cuestión y del grado de error que se pretende obtener. En el caso particular sobre el que se trabaja, interesa no superar el 10% de error. Esto es evaluado mediante el indicador Khat.

Este proceso de aprendizaje permite que se incorpore evidencia al sistema cuando se dispone de ésta, logrando de esta forma una mejora en el comportamiento del clasificador.

El razonamiento evidencial presentado en [Ped95 a] y [Ped95 b] no describe una forma de incorporar nueva evidencia, ni permite evaluar si es necesario hacerlo. Trabaja con un conjunto inicial de muestras de entrenamiento, y un proceso de transformación de bin (ver **capítulo 4, sección 5**) que permite extender dicho conjunto de muestras para lograr una mejor distribución de los datos.

El proceso de aprendizaje tiene por objetivo incorporar evidencia a partir de un conjunto de muestras de entrenamiento, evaluando si el clasificador se comporta con un error por debajo del parámetro de precisión o no. Sólo en el caso que el error sea mayor que este valor se agregan las muestras para modificar la evidencia que se tenía hasta el momento y lograr una mejora en la precisión del mismo. Esto permite extender el conocimiento actual del clasificador, pero dependiendo de un análisis previo que evalúa si la evidencia a incorporar es necesaria para obtener la respuesta del clasificador deseada. Tener errores superiores al 10% implica la necesidad de obtener mayor evidencia para mejorar la respuesta, por lo cual se ejecuta el aprendizaje. El valor del 10% para cota del error puede ser modificado de acuerdo a la precisión deseada para la aplicación con la que se trabaje.

Puede incorporarse de ser necesario nueva evidencia en cualquier momento, y modificar así la base de conocimiento actual, es por ello que surge el nombre de *Dinámico* para el método aquí presentado.

Por otra parte, el otro objetivo de este método es el de realizar un análisis adecuado de las medidas de soporte y plausibilidad obtenidas. Se tienen en cuenta también la cantidad de fuentes intervinientes en la decisión y el peso asignado a las mismas respecto de cada clase. De esta manera, se pretende mejorar la decisión de asignar un objeto a una clase en situaciones donde la evidencia es resultante del aporte de sólo una de las fuentes o es inferior al parámetro de cantidad de fuentes (Rf), y cuando se produce confusión debido a la cercanía de los soportes entre las distintas clases. Es decir, el hecho de optar por la clase con mayor soporte resulta un riesgo cuando no se posee una distancia suficiente entre el soporte de las distintas clases. Por tanto, con todo esto se trata de lograr una decisión más cercana a la realidad.

5.4. Comparación teórica con los métodos convencionales

Los algoritmos de clasificación supervisada presentados en el capítulo anterior se distinguían entre aquellos que no realizaban suposiciones estadísticas (no paramétricos) y aquellos que realizaban tales consideraciones (paramétricos).

Cuando se analizaron los clasificadores supervisados no paramétricos, destacamos que algunos de ellos presentan dificultades al momento de trabajar con clases solapadas, debido a su alta dispersión. Tal es el caso, del clasificador del paralelepípedo (donde no se logra un ajuste con la distribución de los datos cuando los mismos presentan alta covarianza), o los clasificadores basados en regionalización (en los cuales se encuentran problemas con la superposición de clases) [Jen96][Ped93]. En este sentido, el clasificador de razonamiento evidencial dinámico, puede obtener valores de soporte cercanos para clases solapadas y realiza un análisis más específico para tomar la decisión de asignar una clase al objeto desconocido (análisis de ambigüedad, ya explicado en la **sección 2** de este capítulo), por lo que en la mayoría de los casos toma la decisión acertada. No es necesario determinar un rango para cada una de las bandas como realiza el clasificador del paralelepípedo, en donde las regiones determinadas por esos rangos podían no ajustarse a los datos.

El enfoque de redes neuronales presenta ciertas ventajas en cuanto a los inconvenientes de los métodos basados en clasificación supervisada no paramétrica. Se ha mostrado que tanto las redes neuronales como el enfoque de razonamiento evidencial presentan un considerable potencial para el análisis de imágenes a partir de múltiples fuentes y la clasificación de cobertura de suelo [Dug96]. Se pueden destacar algunos aspectos ventajosos que proveen las redes, por un lado, no requieren de la suposición de una cierta distribución para los datos (como en el caso de los clasificadores estadísticos), y por otro, permite integrar diferentes niveles de datos (direccional, nominal, etc.). Sin embargo, el diseño de las redes neuronales se basa en una serie de decisiones que influyen en su posterior rendimiento, por ejemplo, decidir el tipo de conexión entre las neuronas de diferentes capas, así como entre las

neuronas de una misma capa, la forma en que una neurona recibe una entrada y produce una salida, y el peso de cada conexión dentro de la red. Básicamente, se tiene que estudiar la arquitectura de red óptima para lograr un correcto funcionamiento del sistema de clasificación. El método RED requiere de la elección de algunos valores de parámetros, entre ellos, el porcentaje de plausibilidad que se usará como umbral, y los pesos asignados a cada par fuente-clase. El más difícil de establecer puede ser el primero, sin embargo se sabe que valores altos (más del 50%) imponen una restricción mayor para arriesgar por una clase, ya que se pide un alto porcentaje de evidencia no asignada a ninguna clase, y valores bajos permiten arriesgar muchas veces erróneamente por una clase. El valor recomendado en este caso es el 50%. Respecto del peso de los pares fuente-clase, puede resultar sencillo realizar tal asignación si se conoce la aplicación y con qué fuentes se trabaja. En esta tesis, por ejemplo, se trabaja con las diferencias normalizadas de las bandas contiguas y cuatro rasgos de textura de primer orden. Por lo tanto, se sabe que la diferencia normalizada de las bandas correspondientes al rojo y al infrarrojo es significativa a la hora de clasificar vegetación, por lo cual esta fuente de información recibirá mayor peso. En el peor de los casos, se trabaja con un peso igual para todas las posibles combinaciones de fuente-clase, y se estaría considerando sólo el valor de soporte y el índice de cantidad de fuentes aportando evidencia.

Los clasificadores paramétricos, como el de Bayes o el de Máxima Verosimilitud, permiten obtener muy buenos resultados si se dispone de cierta información a priori y si los datos respetan una determinada distribución estadística [Sch92][Tou74]. Para ambos, se requiere asumir una distribución normal de los datos, y además para el clasificador de Bayes es necesario conocer la probabilidad de cada clase. Es decir, a priori se debe contar con cierta información para que el clasificador pueda funcionar de una manera óptima.

Otro problema encontrado es el alto costo computacional de los clasificadores estadísticos, que se incrementa particularmente cuando se tiene un gran número de canales espectrales involucrados o muchas clases entre las que se debe distinguir. Por otra parte, cuando se trabaja con información heurística, una probabilidad de 0.40

para una proposición no necesariamente indica que a su negación se le deba asignar una probabilidad de 0.60 (como en la teoría de Bayes). El razonamiento evidencial y RED realizan una distinción entre "falta de creencia" y "no creencia", por lo que en el caso anterior no considera que si se tiene un soporte de 0.40 para una de las clases, el resto de las clases recibe el 0.60 que queda [Ped93]. El concepto de incertidumbre es importante en la clasificación, especialmente cuando la evidencia es distribuida a partir de una variedad de fuentes diferentes. De esta manera, se provee una representación más adecuada de la situación cuando alguna o toda la información disponible es insuficiente para decidir entre cualquiera de las proposiciones.

Resumiendo, tanto el método de razonamiento evidencial como el razonamiento evidencial dinámico, presentan una serie de ventajas respecto de los clasificadores clásicos mencionados:

- No es necesario presuponer una distribución para los datos.
- La teoría provee una base general y heurística para integrar cuerpos distintos de información a partir de fuentes independientes. De esta forma, pueden incorporarse un alto número de datos y de diferente naturaleza. Incluso es posible incorporar los resultados de una clasificación previa, lo cual puede resultar muy útil para el análisis de la evolución de un área geográfica particular.
- Se permite manejar la evidencia no asignada o incertidumbre. La mayoría de los algoritmos de clasificación no tienen un mecanismo explícito para manejar inconsistencias, errores, o incertidumbre. Esto es particularmente importante con conjuntos de datos dispares, de múltiples fuentes, los cuales pueden poseer niveles diferentes de calidad de información, o instancias de datos perdidos o indefinidos.

Como se puede observar, el método de razonamiento evidencial dinámico propuesto, mantiene las ventajas presentadas en el método de razonamiento evidencial, incorporando una serie de variaciones que modifican en forma positiva el comportamiento del clasificador. La etapa de aprendizaje descrita en la sección

anterior permite la incorporación de nueva evidencia y esto tiende a un notable mejoramiento en los resultados obtenidos (ver **Capítulo 6**). Claro está, que es necesario contar con información confiable sobre la pertenencia de las muestras que se agregan por medio del proceso de aprendizaje a una cierta categoría o clase. Si se dispone de nuevos conjuntos de muestras los mismos pueden mejorar considerablemente los resultados del clasificador.

En cuanto a la regla de decisión propuesta por RED, la misma permite establecer un umbral de seguridad para no arriesgar por una clase a menos que se disponga de una cierta evidencia. Además, la regla trabaja analizando cuál es la clase con mayor soporte, pero no sólo utiliza esta información para la decisión, sino que también estudia la medida de plausibilidad y la distancia entre los soportes de las clases con el fin de saber si otras clases han recibido una evidencia similar. En tal caso, se considera que la clase a asignar es aquella que posee el mayor valor resultante de la combinación del índice de cantidad de fuentes (**ICF**), los pesos de las mismas respecto de la clase, y el soporte. En el **Capítulo 6** esto puede ser visto para la aplicación con la que se trabaja.

5.5. Conclusiones del capítulo

En este capítulo se ha presentado el método de razonamiento evidencial dinámico, sus objetivos, y su descripción. Este método presenta una serie de aspectos positivos: se puede tener como entrada al clasificador un gran número de variables, no es necesario presuponer una determinada distribución estadística, los datos pueden ser de diferente naturaleza y se tiene en cuenta la incertidumbre asociada a los mismos. Todos estos aspectos también son característicos del método de razonamiento evidencial. Sin embargo, la posibilidad de incorporar nueva información para mejorar el rendimiento del clasificador y la decisión final que toma en cuenta no sólo los valores de soporte y plausibilidad, sino también la cantidad de fuentes que aportan evidencia y la importancia de cada fuente para las distintas clases, son características propias del método de razonamiento evidencial dinámico. ◆