

Aprendizaje Autónomo Basado en Mecanismo de Refuerzo aplicado a Sistemas Expertos Legales

Leopoldo Sebastián M. Gómez^{1,2}; Gregorio Perichinsky^{2,3}; Ramón García Martínez^{1,2,4}

¹Programa de Magister en Ingeniería del Software – C.A.P.I.S. – I.T.B.A.

²Programa de Doctorado en Ciencias Informáticas – F.I. – U.N.L.P.

³Laboratorio de Sistemas Operativos y Bases de Datos – F.I. – U.B.A.

⁴Laboratorio de Sistemas Inteligentes – F.I. – U.B.A.

Abstract. Este trabajo propone un modelo de aprendizaje que combina un refuerzo lineal aplicado a las ponderaciones asignadas a los atributos que caracterizan a un caso, con umbrales de decisión dinámicos que son ajustados a partir de casos almacenados en la base de conocimientos. La información incorporada a través de la interacción con el entorno permite mejorar la precisión de las recomendaciones emitidas por sistemas expertos.

Key Words: expert systems, leading cases, machine learning, reinforcement learning, rule-based reasoning.

1. Introducción

Los sistemas expertos son programas de computación que aplican conocimientos sustanciales de áreas específicas de experiencia a la solución de problemas. Estos sistemas son capaces de exhibir comportamiento inteligente por medio de la aplicación precisa de heurísticas.[Wat86]

Uno de los inconvenientes que se plantea, es la dificultad para adaptarlos los cambios en el entorno. Cuando los sistemas expertos mantienen un cuerpo estático de conocimientos, las modificaciones tienen un costo elevado.

2. Descripción del Modelo de Aprendizaje

El modelo de aprendizaje que se expone, permite que un sistema experto pueda adaptar su conocimiento, a medida que conozca los resultados reales de cada caso evaluado. Ver Figura 1.

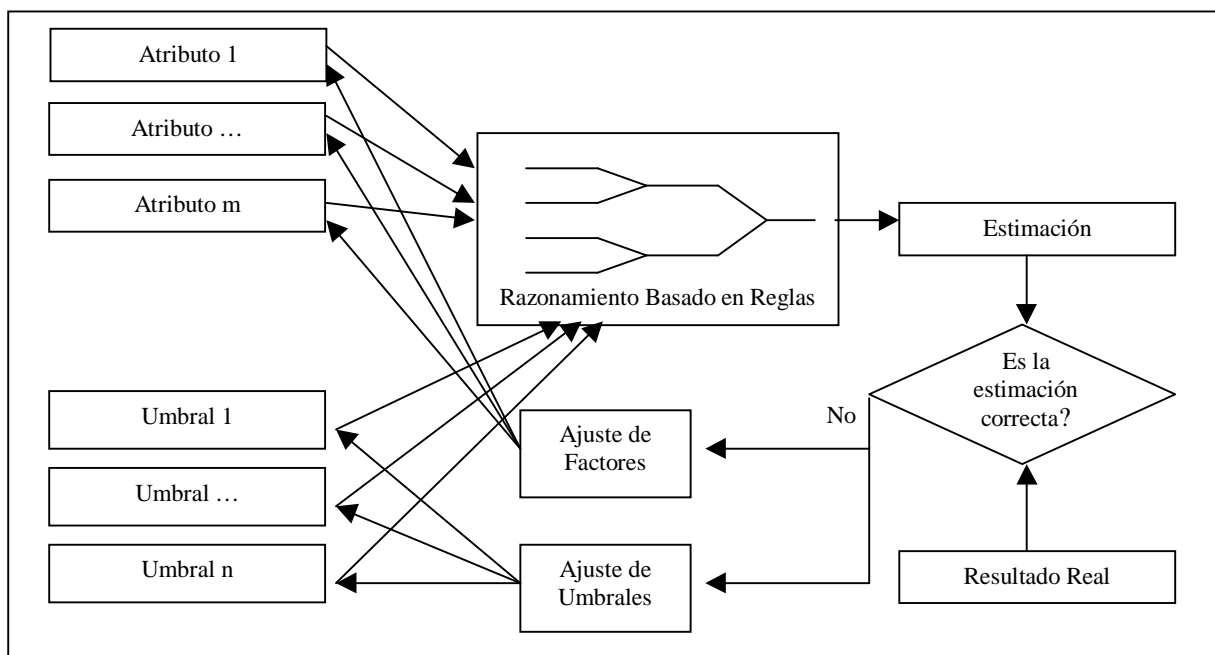


Figura 1. Modelo de aprendizaje para sistemas expertos

Cada uno de los atributos que caracterizan un caso tiene un peso asociado. Mediante razonamiento basado en reglas se obtiene una estimación de la pena está definida por un valor numérico, que luego de ser comparado contra los umbrales de decisión, se traduce en la recomendación, representada por un valor simbólico.

Si una vez conocido el resultado real del caso, no coincide con la recomendación generada por el sistema experto, se aplica un refuerzo a cada uno de los atributos. Dependiendo del valor semántico que posean los atributos, el refuerzo consistirá en un incremento o un decremento de sus ponderaciones, especificado por una tasa de aprendizaje. El comportamiento del sistema converge cuando las representaciones internas asociadas al mismo convergen a valores fijos o que presentan una variación pequeña y lenta. La conducta que se deriva de la fijación de dichos valores puede ser llamada conducta aprendida. [Kae93]

El modelo expuesto se puede plantear como una integración del aprendizaje por refuerzo al proceso de aprendizaje automático de conductas. Existe una vinculación entre la combinación refuerzo-recomendación propuesta en el modelo de aprendizaje con la integración aprendizaje-planificación descripta por García Martínez y Borrajo. [Gar97] y [Gar00]

El proceso de generación de la recomendación y ajuste de los factores y límites, se repetirá hasta lograr predecir correctamente el resultado real del caso. Si bien el proceso de adaptación puede hacer variar las decisiones generadas con anterioridad por el sistema experto, esto tiene sentido, ya que mantiene actualizado sus criterios decisorios.

3. Aplicación del Modelo de Aprendizaje

Para aplicar el modelo de aprendizaje, se ha implementado un sistema experto legal denominado Legal Advisor (LEAD), que tiene como dominio de problema al proceso de la individualización de la pena.

LEAD actúa en forma predictiva, emitiendo una recomendación de la pena para un determinado caso, a partir de sus circunstancias atenuantes y agravantes que puede ser: pena mínima, un tercio de la pena, dos tercios de la pena o pena máxima.

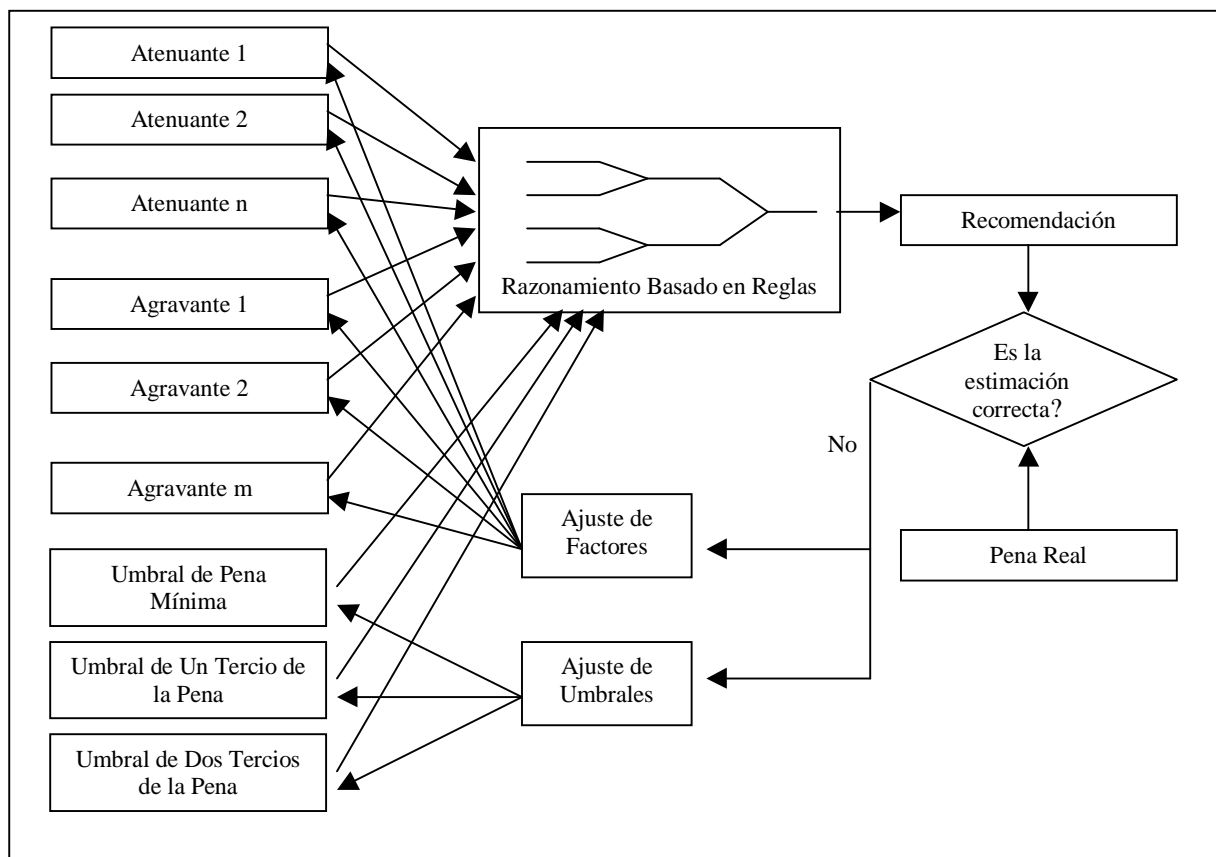


Figura 2. Modelo de aprendizaje aplicado al dominio de la individualización de la pena

La adaptación a los posibles cambios en la ponderación de atributos se logra manteniendo un cuerpo de conocimientos capaz de modificarse en función de los criterios valorativos aplicados en los casos y de las penas reales otorgadas. El modelo de aprendizaje castiga a los factores que intervienen en un caso, ya sea con un aumento o una disminución de sus ponderaciones, cuando la recomendación de la pena propuesta por LEAD no coincide con la pena real otorgada al caso. Ver Figura 2.

En la conceptualización de casos, existe una clasificación a considerar: los casos relevantes - conocidos como “*leading cases*”-, y los casos comunes. Los primeros son aquellos que tienen un efecto profundo sobre las decisiones de todos los casos subsecuentes. En el diseño de LEAD se han considerado las variaciones en los criterios decisorios, introduciendo umbrales que pueden ser ajustados en función de la pena real del caso para obtener mayor precisión en sus recomendaciones. Si el criterio de decisión humana mantiene consistencia en las penas reales, se logrará agrupar a la mayoría de los casos de acuerdo a la pena que les hayan sido otorgadas, pudiendo establecer con precisión los umbrales de decisión.

Dadas las restricciones del dominio del problema entre el intervalo de pena mínima y el de un tercio de la pena, se optó por mantener el Umbral de Pena Mínima estático, ya que se estima que para los casos que no tengan atenuantes ni agravantes, así como también aquellos en que los atenuantes tengan mayor peso que los agravantes la recomendación será Pena Mínima. Sin embargo, los umbrales de Un Tercio y Dos Tercios de la Pena, se actualizan a medida que se incorporan los “*leading cases*” de cada intervalo de decisión a la base de conocimientos de LEAD. Ver Figura 3.

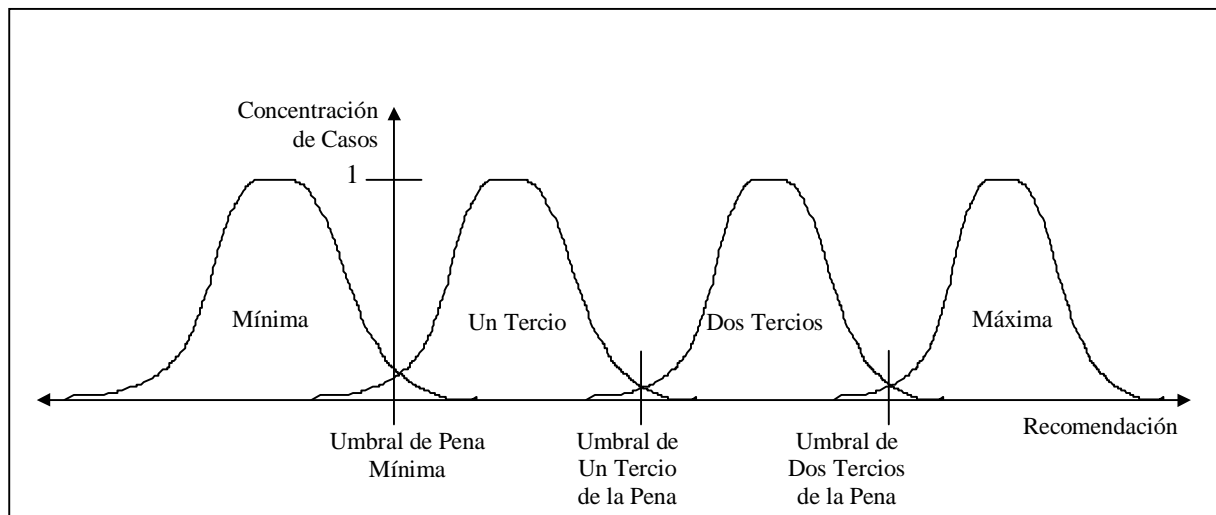


Figura 3. Concentración de casos en función de las penas

4. Conclusiones

El modelo de aprendizaje descrito puede ser incorporado a sistemas expertos para dotarlos de capacidades adaptativas. Por otra parte, este modelo tiene una aplicación directa sobre sistemas expertos predictivos que utilicen una ponderación cuantitativa para indicar la influencia de cada atributo sobre la conclusión final.

La integración del modelo de aprendizaje en un sistema experto deriva en dos ventajas significativas: se reduce la necesidad de reprogramar el cuerpo de conocimientos y ayuda a mantener la eficacia de sus recomendaciones, lo que garantiza su utilidad a lo largo del tiempo.

5. Referencias

- [Gar97] García Martínez, R. y Borrajo, D. “*Planning, Learning and Executing in Autonomous Systems*”. Lecture Notes in Artificial Intelligence. Nber. 1348(Ed. Sam Steel & Rachid Alami) Páginas 208-210. Springer-Verlag. 1997.
- [Gar00] García Martínez, R. y Borrajo, D. “*An Integrated Approach of Learning, Planning and Executing*”. Journal of Intelligent and Robotic Systems. Volumen 29, Número 1, Páginas 47-78. Kluwer Academic Press. 2000
- [Kae93] Kaelbling, L. “*Learning in Embedded Systems*”. MIT Press. 1993.
- [Wat86] Waterman, Donald A. “*A Guide to Expert Systems*”, Reading, MA: Addison-Wesley, 1986.