

# CARACTERIZACIÓN DE PROBLEMAS DE APRENDIZAJE BASADA EN EXPLOTACIÓN DE INFORMACIÓN

Elizabeth Jiménez Rey, Darío Rodríguez, Paola Britos, Ramón García-Martínez

Programa de Maestría en Tecnología Informática Aplicada a la Educación. Facultad de Informática. UNLP  
Laboratorio de Sistemas Operativos y Base de Datos. Facultad de Ingeniería. UBA  
Centro de Ingeniería de Software Ingeniería del Conocimiento. Escuela de Postgrado. ITBA  
Laboratorio de Sistemas Inteligentes. Facultad de Ingeniería. UBA  
Área Ingeniería del Software. Licenciatura Sistemas. UNLa  
Área Ingeniería del Software. Unidad Académica Río Gallegos. UNPA

ejimenezrey@yahoo.com.ar, {drodrigu,pbritos}@itba.edu.ar, rgarciamar@fi.uba.ar

## CONTEXTO

La propuesta articula líneas de trabajo del Proyecto “Aplicaciones de Explotación de Información basada en Sistemas Inteligentes”, con financiamiento de la Secretaría de Ciencia y Técnica de la Universidad de Buenos Aires (UBACYT 2008-2010 código I012) y acreditado por Resolución Rector-UBA N° 576/08 con radicación en el Laboratorio de Sistemas Inteligentes de la Facultad de Ingeniería de la Universidad de Buenos Aires

## RESUMEN

Se presenta una línea de investigación que se focaliza en la aplicación de herramientas de Minería de Datos basada en Sistemas Inteligentes para la identificación y predicción de problemas de aprendizaje de los estudiantes (y sus causas).

El interés principal radica en la exploración de la posibilidad de construir un proceso que permita al Docente, por una parte, revisar la precedencia de los aprendizajes significativos considerados en el diseño del currículo y, por otra, reflexionar sobre las estrategias de su práctica docente para promover su mejoramiento.

Se aborda el estudio de los diversos métodos de la Minería de Datos: algoritmos de inducción, algoritmos genéticos, redes neuronales y redes bayesianas, para establecer el método que resultará más confiable para resolver el problema.

Actualmente, el trabajo de investigación está centrado en el estudio de los conceptos no comprendidos por los estudiantes de dos cursos de la asignatura Computación de carreras no informáticas de la Facultad de Ingeniería de la Universidad de Buenos Aires.

**Palabras clave:** *Explotación de Información aplicada en Educación, Problemas de Aprendizaje en Educación.*

## 1. INTRODUCCIÓN

En esta línea de investigación, el interés se centra en permitir a los Docentes la revisión de sus estrategias de enseñanza para favorecer en los estudiantes el

desarrollo de la cognición, en su más amplio sentido: la comprensión y la capacidad para pensar y conocer cada vez mejor [Litwin, 2008]. En este escenario, se explorará la posibilidad de construir un proceso que permita al Docente revisar la precedencia de los aprendizajes significativos [Joyce y Weil, 2002] de los estudiantes considerados en el diseño del currículo. La solución propuesta es una segmentación de los procesos intuitivos del Docente (de construcción artesanal y fuertemente ligados a su experiencia) que se estructura en un proceso de varias etapas: 1) Construcción del Mapa de Aprendizajes Significativos, que se articulará en concepto-atributo-valor [García Martínez y Britos, 2004] a determinar su existencia en los instrumentos de evaluación. 2) Explotación de Información sobre la base de evaluación. 3) Ratificación o Rectificación del Mapa de Aprendizajes Significativos.

Se estudia actualmente los casos correspondientes a dos cursos de la asignatura Computación de carreras no informáticas de la Facultad de Ingeniería de la Universidad de Buenos Aires.

La asignatura Computación es formativa, básica y obligatoria para alumnos de todas las carreras de Ingeniería, excepto para aquellos que cursan Ingeniería Electrónica e Ingeniería en Informática. Es cuatrimestral y tiene una carga horaria de cuatro horas por semana (sesenta y cuatro horas por cuatrimestre). La única asignatura correlativa posterior es Análisis Numérico.

Los alumnos pueden cursar la asignatura Computación en distintas etapas de avance en sus planes de estudio pero casi todos los alumnos la cursan en los primeros años de su carrera y tienen una edad comprendida entre 18 y 20 años. Algunos pocos alumnos provienen de escuelas técnicas y tienen conocimiento y experiencia previa en programación pero la mayoría de los alumnos nunca programó. La algoritmia representa para los alumnos un nuevo paradigma para resolver problemas y les produce un fuerte impacto en su predisposición al aprendizaje que, en muchos casos, se traduce en falta de motivación o en rechazo.

Son los objetivos *generales* de enseñanza de la asignatura:

- Brindar al alumno una visión global de la Computación, de manera que comprenda conceptos y técnicas de la disciplina que en su futuro profesional lo habilitarían para interactuar en forma interdisciplinaria con pares y profesionales en informática sin problemas de comunicación.
- Lograr que el alumno se compenetre con las tecnologías y herramientas fundamentales de la Computación, de manera que aprenda a usar la computadora como herramienta de trabajo, conociendo su precisión, capacidad y limitaciones.

Son los objetivos *específicos* de enseñanza de la asignatura:

- Concientizar al alumno de la importancia de la Algoritmia como paradigma de resolución de problemas y de la Programación como práctica y ejercitación en la resolución de problemas, como promotores del desarrollo de la capacidad de abstracción, es decir, la capacidad de relacionar esquemas de solución con la resolución de problemas algorítmicos, con hincapié en el método científico para lograr ese objetivo.
- Enseñar al alumno de Carreras de Ingeniería el Análisis, la Sistematización, Programación y Procesamiento de distintos problemas de tipo técnico-científicos, a fin de que dichos conocimientos le resulten de utilidad ya sea en el desarrollo de la carrera como así también en su actividad profesional.
- Sustentar la estrategia de enseñanza y de aprendizaje de los conceptos y procedimientos en los recursos tecnológicos, como mediadores de los procesos y las habilidades cognitivas a inducir en el alumno.

Son los objetivos de aprendizaje de la asignatura:

- Conocimiento Declarativo (SABER que el alumno debe adquirir): capacidad de resolución de problemas con la computadora. El alumno deberá conocer QUÉ es el Hardware (computadora) y QUÉ es el Software (programas).
- Conocimiento Procedimental (SABER HACER que el alumno debe desarrollar): habilidad para encontrar la Solución (representada por un programa) a un Problema (representado por un enunciado), es decir, CÓMO construir un programa solución a partir de un enunciado problema.
- Conocimiento Actitudinal (SABER ESTAR): dominio de actitudes de compromiso, responsabilidad, solidaridad y respeto del estudiante en el proceso de construcción de su conocimiento. La interacción continua del alumno con sus pares, con el Docente y con el material de estudio, favorecerá la comprensión, la reflexión, la acción y la comunicación, aspectos claves que el alumno deberá poner en juego en su proceso de aprendizaje. La estrategia de enseñanza y de aprendizaje de la asignatura Computación se focaliza en la adopción del proceso de resolución de problemas de Polya

[Polya, 1945] como sustento para la construcción de programas, es decir, la resolución de problemas con la computadora. Se representa dicho proceso, en el contexto de la creación de programas, por medio de un Mapa Conceptual [Novack, Gowin, y Otero, 2002]. El Mapa Conceptual, utilizado iterativamente durante el desarrollo del curso, a manera de ciclo, es el instrumento que, por una parte, permite adoptar el método científico como estrategia para el desarrollo del contenido (el qué y el cómo enseñar y aprender) en forma evolutiva e incremental [Jiménez Rey, 2005] y, por otra parte, es la herramienta estratégica educativa que permite implementar el proceso de aprendizaje humano (modelo de la memoria de dos almacenes) [Shunk, 1997], facilitando al estudiante la comprensión del conocimiento conceptual, procedimental y operativo.

La concepción del Docente es que un programa se construye siguiendo un proceso evolutivo (Modelo de Polya de las Cuatro Fases para la Resolución de Problemas) que abarca cuatro fases:

**Análisis:** El estudiante debe comprender en qué consiste el problema a resolver con la construcción del programa. Debe ser capaz de interpretar el enunciado del problema a resolver con la computadora, es decir, identificar y definir el objetivo del programa.

**Diseño:** El estudiante debe idear un plan para construir el programa. Debe ser capaz de entender cuál es la naturaleza del problema a resolver, es decir, identificar los recursos necesarios a utilizar y el procedimiento de transformación de esos recursos que conduce al alcance del objetivo.

**Codificación:** El estudiante debe implementar el plan para construir el programa, o sea, representar los recursos y el procedimiento necesarios en un lenguaje de programación para que la computadora pueda resolver el problema. Debe ser capaz de comprender la semántica del lenguaje Pascal y expresar el algoritmo en la sintaxis correspondiente, es decir, expresar la solución del problema como programa concreto verificable.

**Evaluación:** El estudiante debe comprobar que el programa diseñado para resolver el problema especificado en el enunciado es una solución ejecutable y correcta. Debe ser capaz de diseñar los juegos de datos de prueba representativos del problema para que la ejecución del programa conduzca al resultado deseado en todas las situaciones posibles del problema, independientemente de los datos particulares de una ejecución (eliminar y depurar errores). La explotación de información ha sido señalada como una manera efectiva de descubrir nuevo conocimiento desde conjuntos de datos de procesos educacionales, datos generados por sistemas de aprendizaje o experimentos de aprendizaje, así como la información descubierta puede ser usada para probar la personalización y la adaptación [Britos, 2008]. Entre los problemas interesantes que la

explotación de información puede ayudar a resolver: determinación de cuáles son los estilos o estrategias de aprendizaje comunes [Britos *et al.*, 2008a], predicción del conocimiento e intereses de un usuario basado en su comportamiento previo, particionamiento de un grupo heterogéneo de usuarios en clusters homogéneos o detección de conceptos no comprendidos en procesos de aprendizaje [Jiménez Rey *et al.*, 2008; Britos *et al.*, 2008c].

Una de las técnicas más comunes de explotación de información son los árboles de decisión (TDIDT), usados para descubrir conocimiento en forma de reglas, las cuales constituyen un modelo que representa el dominio del conocimiento subyacente a los ejemplos disponibles sobre el mismo. Una red bayesiana es un gráfico acíclico direccionado en el cual cada nodo representa una variable y cada arco representa una dependencia probabilística, la cual especifica la probabilidad condicional de cada variable dada a sus padres; la variable a la cual el arco apunta es dependiente (causa – efecto) de la variable en el origen de ésta [Britos *et al.*, 2008b].

## 2. LÍNEAS DE INVESTIGACIÓN y DESARROLLO

Las líneas de investigación abarcan los siguientes ejes temáticos:

- Enseñanza de Computación en Carreras de Ingeniería
- Enseñanza para la Comprensión
- Aprendizaje basado en Problemas
- Redes Conceptuales
- Teoría Triárquica de la Inteligencia
- Análisis del Error
- Nuevas Tecnologías de la Información y de la Comunicación en la Enseñanza de Algoritmia y Programación
- Educación Basada en Web
- Aplicaciones de Explotación de Información basada en Sistemas Inteligentes a Educación

## 3. RESULTADOS OBTENIDOS/ESPERADOS

Se ha identificado y clasificado, una lista de componentes de conceptos no comprendidos del modelo del estudiante a ser evaluado, teniendo en consideración tres aspectos fundamentales en la creación de programas: Metodología de Desarrollo, Funcionalidad del Programa y Calidad del Diseño.

Para descubrir conceptos mal comprendidos de programación de los estudiantes, se focalizó la investigación en el uso de herramientas de bases de datos de sistemas inteligentes: inducción de reglas por algoritmos TDIDT y redes bayesianas. Estas reglas se utilizaron en un proceso de descubrimiento del conocimiento en tres pasos:

- Construcción de una base de datos sobre una caracterización estándar de cada estudiante y su estilo y conceptos mal aprendidos de programación.

- Descubrimiento de reglas (mediante algoritmos TDIDT) las cuales establecieron una relación entre conceptos no comprendidos de programación y sus posibles causas.

- Descubrimiento del peso que cada causa tiene sobre cada concepto mal aprendido (mediante redes bayesianas).

En la actualidad, con el objetivo a preparar una metodología escalable, se está revisando los modelos de representación de precedencia de aprendizajes significativos, previéndose realizar un “tuning” del procedimiento experimental seguido para las pruebas de concepto.

## 4. FORMACIÓN DE RECURSOS HUMANOS

En el marco de este proyecto se está desarrollando una tesis de maestría y dos trabajos de especialidad radicadas en el Laboratorio de Sistemas Inteligentes de la Facultad de Ingeniería de la Universidad de Buenos Aires.

## 5. BIBLIOGRAFÍA

- Britos, P. (2008). *Procesos de Explotación de Información Basados en Sistemas Inteligentes*. Tesis de Doctorado en Ciencias Informáticas. Facultad de Informática. UNLP.
- Britos, P., Cataldi, Z., Sierra, E., García-Martínez, R. (2008a). *Pedagogical Protocols Selection Automatic Assistance*. LNAI, 5027: 331-336.
- Britos, P., Felgaer, P., Garcia-Martinez, R. (2008b). *Bayesian Networks Optimization Based on Induction Learning Techniques*. IFIP Series, 276: 439-443.
- Britos, P., Jiménez Rey, E., Rodríguez, D., García-Martínez, R. (2008c). *Work in Progress: Programming Misunderstandings Discovering Process Based On Intelligent Data Mining Tools*. Proceedings 38th ASEE/IEEE Frontiers in Education Conference. Session F4H: Assessing and Understanding Student Learning. ISBN 978-1-4244-1970-8.
- García Martínez, R. y Britos, P. (2004). *Ingeniería de Sistemas Expertos*. Editorial Nueva Librería. ISBN 987-1104-15-4.
- Jiménez Rey, E. (2005). *Un Enfoque Procedimental para la Enseñanza de Computación en Carreras de Ingeniería*. Proceedings JEITICs 2005: 35-39.
- Jiménez Rey, E., Rodríguez, D., Britos, P., García-Martínez, R. (2008). *Identificación de Problemas de Aprendizaje de Programación con Explotación de Información*. Proceedings del XIV CACIC 2008. Artículo 1881. ISBN 978-987-24611-0-2.
- Joyce, B. y Weil, M. (2002). *Modelos de Enseñanza*. Gedisa. ISBN 84-7432-780-6.
- Litwin, E. (2008). *El Oficio de Enseñar, Condiciones y Contextos*. Paidós. ISBN 978-950-12-1513-7.
- Polya, G. (1945). *How to Solve It: A New Aspect of Mathematical Method*. Princeton. ISBN 0-691-08097-6.