

Proceso de Identificación de Errores de Apropiación de Conceptos Basado en Explotación de Información

Saavedra-Martínez, P., Pollo-Cattaneo, F., Rodríguez, D., Britos, P., García-Martínez, R.

Grupo de Investigación en
Sistemas de Información.
Departamento de Desarrollo
Productivo y Tecnológico.
Universidad Nacional de Lanús.
rgarcia@unla.edu.ar

Programa de Maestría en
Tecnología Informática Aplicada a
la Educación. Facultad de
Informática. Universidad Nacional
de La Plata
psaavedra_martinez@hotmail.com

Grupo de Estudio en Metodologías
de Ingeniería de Software. Escuela
de Posgrado. Facultad Regional
Buenos Aires. Universidad
Tecnológica Nacional.
fpollo@posgrado.frba.utn.edu.ar

Grupo de Investigación en
Explotación de Información
Sede Andina
Universidad Nacional de Río
Negro
paobritos@gmail.com

Resumen

Este trabajo presenta resultados de un trabajo de investigación en progreso que busca definir un proceso de identificación de errores de apropiación de conceptos, utilizando herramientas de explotación de información. El proceso propuesto está estructurado en cinco pasos: conceptualización del dominio, identificación del subdominio de análisis, preparación de los datos, explotación de información e interpretación de los resultados. Se presenta una prueba de concepto de la utilización del proceso en el dominio de la asignatura Análisis de Sistemas.

Palabras claves: Errores de Apropiación de Conceptos, Explotación de Información

1. Introducción

Para Spotts y Bowman [1995] el aprendizaje es definido como un cambio estable en la conducta. Antes de aprender un alumno no hace algo y, después del aprendizaje, es capaz de hacerlo. En este contexto, la evaluación es la herramienta que utilizan los docentes que les permite observar los avances en el desarrollo de las habilidades de conducta, los objetivos del programa y detectar los errores frecuentes, evaluando la eficacia del programa de la materia propuesto. Sin embargo las evaluaciones, por sí mismas, no permiten al docente diagnosticar las dificultades de aprendizaje de los alumnos y sus causas. Por ello se introduce, en este contexto, el concepto de explotación de información.

Se define [Larose, 2005] el término explotación de información (Data Mining) como el proceso de descubrir nuevas correlaciones, patrones y tendencias, utilizando grandes cantidades de datos almacenados en repositorios, aplicando tecnologías de reconocimiento de patrones, así como herramientas matemáticas y estadísticas. La explotación de información es un elemento fundamental de un proceso más amplio que tiene como objetivo el descubrimiento de conocimiento en grandes bases de datos [Fayyad *et al.*, 1996; Britos *et al.*, 2005], en inglés “Knowledge Discovery in Databases” (KDD).

El objetivo principal del proyecto radica en la obtención de una metodología que permita al docente: (1) identificar los errores de aprendizaje de los alumnos en instancias evaluativas y (2) diagramar los conceptos enseñados en pos de minimizar, en tanto sea posible, dichos errores. Para poder cumplir con estos objetivos, el docente debe ser capaz, en primera instancia, de determinar el mapa conceptual y de precedencias de los temas comprendidos.

El proceso citado contempla: desde la generación de un modelo conceptual de los temas enseñados (de tratamiento intuitivo quizás en el dictado del curso) hasta la conceptualización del proceso o modelo, pasando por la preparación de datos, la explotación de los mismos y la posterior obtención de reglas y patrones, que permitan la evaluación e interpretación de los resultados que llevan a la obtención del conocimiento.

2. Proceso de Identificación de Errores de Apropiación Propuesto

El proceso de identificación de errores de apropiación de conceptos en los estudiantes, se basa en trabajos previos de los autores [Britos et al., 2008; Jiménez Rey et al., 2008; 2009; 2010]. Se utilizan los procesos de explotación de información [Britos, 2008; Britos y Garcia-Martinez, 2009]: descubrimiento de reglas de comportamiento y ponderación de interdependencia de atributos. Con base en estos procedimientos se define el siguiente proceso de cinco pasos para identificar errores de apropiación de conceptos:

Paso 1: Conceptualización del Dominio

- Subpaso 1.1: Seleccionar los conceptos enseñados que se consideran importantes para la determinación del aprendizaje del individuo.
- Subpaso 1.2: Construir el Glosario de Términos.
- Subpaso 1.3: Construir el mapa de dependencia o precedencia conceptual.

Paso 2: Identificación del Subdominio de Análisis

- Subpaso 2.1: Identificar el submapa candidato para el análisis.
- Subpaso 2.2: Refinar los conceptos del submapa candidato en subconceptos.
- Subpaso 2.3: Identificar los atributos de evaluación vinculados al submapa candidato en subconceptos.

Paso 3: Preparación de los Datos

- Subpaso 3.1: Identificar los valores de los atributos de evaluación, vinculados a los subconceptos del submapa candidato en las instancias de evaluación individual.
- Subpaso 3.2: Construir la Base de Datos, usando como campos los atributos de evaluación y, como instancias, los valores de estos atributos para cada sujeto de la población en estudio.

Paso 4: Explotación de Información

- Subpaso 4.1: Aplicar el proceso Descubrimiento de Reglas de Comportamiento [Britos, 2008], tomando como clase el atributo que se encuentra en la raíz del subárbol asociado al del submapa candidato
- Subpaso 4.2: Aplicar el proceso Ponderación de Interdependencia de Atributos [Britos, 2008], tomando a los atributos hoja de las reglas obtenidas en el Subpaso 4.1 y el atributo clase.

Paso 5: Interpretar los resultados.

3. Prueba de Concepto

En esta sección se presenta una prueba de concepto del proceso propuesto. Se delimita el dominio de conocimiento (sección 3.1) y se presenta un caso de estudio que resulta de aplicar el proceso propuesto a dicho dominio (sección 3.2).

3.1. Delimitación del Dominio

Para este trabajo hemos tomado como caso de estudio los conceptos enseñados en la asignatura Análisis de Sistemas (2do año de la carrera Ingeniería en Sistemas de Información de la Universidad Tecnológica Nacional Facultad Regional Buenos Aires (UTN - FRBA); que se focaliza en la enseñanza de las herramientas de: Diagrama de Contexto, Tabla de Eventos, DFD (Diagrama de Flujo de Datos) y Diccionario de Datos.

Los objetivos básicos de la asignatura son:

- Elaborar modelos conceptuales de un sistema de información.
- Conocer las distintas etapas del proceso de desarrollo de sistemas de información.
- Modelar las características intrínsecas de los sistemas de información.
- Conocer y aplicar las metodologías, modelos, técnicas y lenguajes de la etapa de análisis.

- Seleccionar adecuadamente los modelos que mejor se adapten para dar soluciones a los problemas de información.
- Conocer y aplicar los elementos que componen la ingeniería de requerimientos.
- Documentar el proceso de análisis de sistemas de información

No es el contenido del aprendizaje lo que más atrae en el proceso de aprendizaje, sino las actividades que se implementan. El aprendizaje se realiza en el proceso de su ejecución. [Aebli, Hans]

Los alumnos necesitan programas de enseñanza y, los docentes, diseñan planes de lecciones donde se describe el “saber” que el alumno debe haberse apropiado al final de la lección.

Existe un aprendizaje estructural donde los procedimientos de la motivación para el aprendizaje pueden analizarse en el marco de cuatro funciones [Aebli, 1991]:

3.1.1. Motivación para la Construcción por Solución de Problemas

Si aprender significa construir a partir de elementos el nuevo pensamiento donde la clave del proceso es la concatenación de los mismos, la motivación surge del problema o situación que se plantea. Allí donde se plantea un problema, se está próximo a intentar una primera solución, el fracaso relativo proporciona el punto de partida para otras posibles soluciones, lo importante es que el alumno debe captar que avanza, debe notar un progreso en su comprensión.

3.1.2. Motivación del Aprendizaje en la Elaboración

La elaboración hace móviles y aplicables las operaciones y conceptos. Por ejemplo, si se pide a un alumno que elabore un ejemplo sobre un tema dado, para lograr el resultado esperado, debe haber analizado ese tema desde puntos de vista diferentes hasta estar completamente familiarizado con él. Con la elaboración, se motiva a los alumnos al dominio flexible del concepto y la operación, al elaborar algo nuevo, el alumno puede medir

el grado de avance en la comprensión del tema.

3.1.3. Motivación del Aprendizaje en el Ejercicio

En el ejercicio lo que se busca es que un procedimiento se realice inmediata, segura y rápidamente. La motivación implica plantear determinados “retos” a los alumnos que les permita ver claramente el avance en el tema. El tiempo empleado para la ejecución de un procedimiento conocido, llegar a un resultado correcto previamente conocido, puede ser una herramienta importante para que el alumno haga perceptible su progreso. Promover la autoevaluación, también permite dirigir la atención del alumno que actúa, ya que aprende a observarse, a observar su manera de trabajar, a tener en cuenta las características del proceso y su propio funcionamiento.

3.1.4. Motivación del Aprendizaje en la Aplicación

Cuando se ha avanzado lo suficiente como para que se pueda aplicar un concepto o procedimiento nuevo en otras situaciones, el problema de la motivación se simplifica: una situación de aplicación es de nuevo una situación problema, que incita a su dominio, aún cuando en el aula se puedan plantear diversas situaciones de aplicación, cuanto más natural y vivencial sea la situación, más se interesarán los alumnos en ella, siendo concientes de su capacidad para manejar la nueva situación.

3.2. Caso de Estudio

En esta sección se aplican los pasos y subpasos del proceso propuesto a una base de datos de 201 registros.

3.2.1. Subpaso 1.1

Se seleccionan los siguientes conceptos del dominio: Proceso, Entrada, Salida, Entidades Externa, Diagrama de Contexto (DC), Tabla

de Eventos (TE), Diagrama de Flujo de Datos (DFD).

3.2.2. Subpaso 1.2

Los términos descriptivos del caso de estudio quedan planteados en el siguiente glosario de términos:

DC (Diagrama de Contexto): diagrama que representa el intercambio de información entre el sistema de información analizado y las entidades externas.

DFD (Diagrama de Flujo de Datos): diagrama que el modelo lógico de funcionalidad del sistema de información analizado. Muestra los procesos que forman parte del sistema incluyendo las entradas y salidas generadas por cada uno de ellos. También permite visualizar la comunicación (en cuanto a información) que hay entre los procesos.

Entidades Externas: entes u objetos que no pertenecen al sistema de información que tenemos que modelar, pero que tienen relación con el mismo, ya sea generando o recibiendo información de los procesos del sistema.

Entrada: datos que sirven de materia prima para la ejecución de un proceso.

Proceso: conjunto de actividades o funciones que transforman información.

Salida: resultado o respuesta que se obtiene de la ejecución de un proceso.

TE (Tabla de Eventos): diagrama que muestra las funciones o procesos que forman parte del alcance del sistema de información analizado.

3.2.3. Subpaso 1.3 y Subpaso 2.1

El mapa de dependencia o precedencia conceptual construido se presenta en la figura 1. Los conceptos sombreados corresponden al submapa candidato para el análisis.

3.2.4. Subpaso 2.2

El refinamiento de los conceptos del submapa candidato en subconceptos se presenta en la Figura 2.

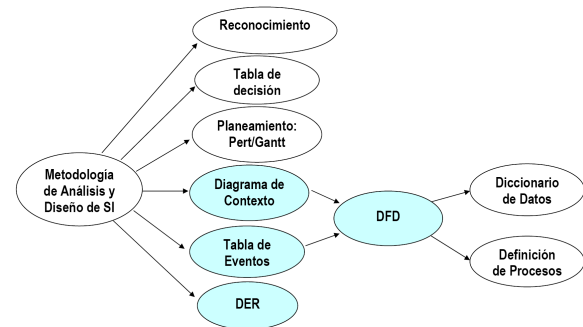


Fig. 1. Mapa de dependencia y submapa candidato para el análisis

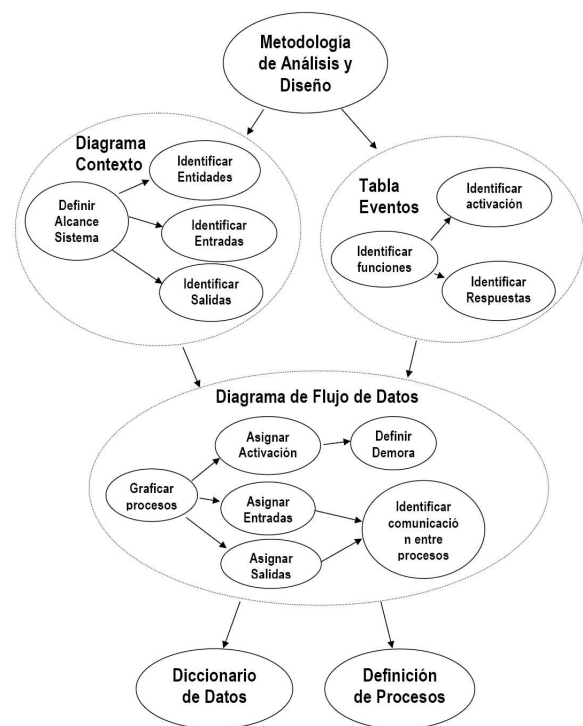


Fig. 2. Refinamiento en subconceptos de los conceptos del submapa candidato.

3.2.5. Subpaso 2.3

Los atributos de evaluación vinculados al submapa candidato en subconceptos que se identifican son (si el alumno es capaz de): definir alcance sistema, identificar entidades, identificar entradas, identificar salidas, identificar funciones, identificar activación, identificar respuestas, graficar procesos, asignar activación, asignar entradas, asignar salidas, definir demoras e identificar comunicación entre procesos.

3.2.6. Subpaso 3.1

Los valores de los atributos de evaluación vinculados a los subconceptos del submapa candidato, en las instancias de evaluación individual, son “identificación de entradas y salidas”. En los mismos se evalúa si el alumno puede determinar cuáles son los flujos que actúan como entrada del sistema que, a su vez, le permitan generar respuestas (salidas). Para los otros atributos los valores que se pueden asignar son: Bien (B), Regular (R) y Mal (M).

3.2.7. Subpaso 3.2

Se desarrolló una Base de Datos con 101 registros utilizando como campos los atributos de evaluación y, como instancias, los valores de estos atributos, para cada sujeto de la población en estudio

3.2.8. Subpaso 4.1

Se aplicó el proceso Descubrimiento de Reglas de Comportamiento, tomando como clase el atributo que se encuentra en la raíz del subárbol asociado al del submapa candidato, dando como resultado, conjuntos de reglas. Estas proporcionan piezas de interés para la revisión docente de los procesos de enseñanza-aprendizaje. Las reglas resultantes se presentan en las Tablas 1, 2 y 3.

3.2.9. Subpaso 4.2

Se aplicó el proceso Ponderación de Interdependencia de Atributos tomando a los atributos hoja de las reglas obtenidas en el Subpaso 4.1 y el atributo clase. Se identificaron las siguientes preguntas para la aplicación del proceso y el correspondiente resultado de la aplicación del proceso de explotación de información:

[a] ¿Cuál de los siguientes atributos: “Identifica entradas y salidas = B”, “Identifica entidades externas = B”, “Identifica entradas y salidas = R” incide más sobre el atributo “DC= B”?

Tabla 1. Reglas sobre Diagrama de Contexto (DC)

REGLA 1	
SI	Identifica entradas y salidas = B
ENTONCES	DC = B [0,9667]
REGLA 2	
SI	Identifica entidades externas = B
Y	Identifica entradas y salidas = R
ENTONCES	DC = B [0,7200]
REGLA 3	
SI	Identifica entidades externas = M
Y	Identifica entradas y salidas = R
ENTONCES	DC = M [1,0000]
REGLA 4	
SI	Identifica entidades externas = R
Y	Identifica entradas y salidas = R
ENTONCES	DC = R [0,6364]
REGLA 5	
SI	Identifica entradas y salidas = M
ENTONCES	DC = M [0,8788]

Tabla 2. Reglas sobre Tabla de Eventos (TE)

REGLA 6	
SI	Puede identificar funciones = R
ENTONCES	Entonces TE = M [0,5000]
REGLA 7	
SI	Identifica las respuestas = M
Y	Puede identificar funciones = B
ENTONCES	TE = M [0,8000]
REGLA 8	
SI	Identifica las respuestas = R
Y	Puede identificar funciones = B
ENTONCES	Entonces TE = B [0,6129]
REGLA 9	
SI	Si Identifica las respuestas = B
Y	Puede identificar funciones = B
ENTONCES	Entonces TE = B [0,9000]
REGLA 10	
SI	Puede identificar funciones = M
ENTONCES	Entonces TE = M [0,9677]

Tabla 3. Reglas sobre Diagrama de Flujos de Datos (DFD)

REGLA 11	
SI	Asigna las entradas y salidas a cada proceso = M
ENTONCES	DFD = M [1,0000]
REGLA 12	
SI	Define las demoras = M
Y	Asigna las entradas y salidas a cada proceso = R
ENTONCES	DFD = M [1,0000]
REGLA 13	
SI	Define las demoras = B
Y	Asigna las entradas y salidas a cada proceso = R
ENTONCES	DFD = R [0,4000]
REGLA 14	
SI	Define las demoras = R
Y	Asigna las entradas y salidas a cada proceso = R
ENTONCES	DFD = M [0,6250]
REGLA 15	
SI	Define las demoras = M
Y	Asigna las entradas y salidas a cada proceso = B
ENTONCES	DFD = R [1,0000]
REGLA 16	
SI	Define las demoras = R
Y	Asigna las entradas y salidas a cada proceso = B
ENTONCES	DFD = B [0,5714]
REGLA 17	
SI	Identifica la comunicación entre procesos = B
Y	Define las demoras = B
Y	Asigna las entradas y salidas a cada proceso = B
ENTONCES	DFD = B [1,0000]
REGLA 18	
SI	Identifica la comunicación entre procesos = R
Y	Define las demoras = B
Y	Asigna las entradas y salidas a cada proceso = B
ENTONCES	Entonces DFD = B [0,8889]

REGLA 19

SI	Identifica la comunicación entre procesos = M
Y	Define las demoras = B
	Asigna las entradas y salidas a cada proceso = B
ENTONCES	DFD = R [0,8000]

Resultado: El atributo que incide más sobre el atributo "DC=B" es "Identifica entidades externas = B", en un 93,88%", le sigue "Identificar entradas y salidas = B" en un 59,18%, finalmente "Identificar entradas y salidas = R" en un 40,82%

[b] ¿Cuál de los siguientes atributos: "Identifica entidades externas = M", "Identifica entradas y salidas = R", "Identifica entradas y salidas = M" incide más sobre el atributo "DC = M"?

Resultado: El atributo que incide sobre "DC=M", es "Identifica entradas y salidas" en un 80,56% si es M, 19,44% si es R; en cuanto a "Identifica entidades externas" en un 50% si es M.

[c] ¿Cuál de los siguientes atributos: "Identifica entidades externas = R", "Identifica entradas y salidas = R" incide más sobre el atributo "DC = R"?

Resultado: El atributo que incide sobre "DC=R", es "Identifica entradas y salidas" en un 68,75% si es R; en cuanto a "Identifica entidades externas" en un 56,25% si es R.

[d] ¿Cuál de los siguientes atributos: "Puede identificar funciones = R", "Identifica las respuestas = M", "Puede identificar funciones = B", "Puede identificar funciones = M" incide más sobre el atributo "TE = M"?

Resultado: El atributo que incide sobre "TE=M", es "Identifica las respuestas = M" en un 79,49%, sigue "Puede identificar funciones = M" en un 76,42%, posteriormente "Puede identificar funciones = B" en 19,75 un %, finalmente "Puede identificar funciones = R" en un 5,31%.

[e] ¿Cuál de los siguientes atributos: “Identifica las respuestas = R”, “Puede identificar funciones = B”, “Identifica las respuestas = B” incide más sobre el atributo “TE = B”?

Resultado: El atributo que incide sobre “TE = B” mayormente es “Puede identificar funciones = B” en un 97,92%, sigue “Identifica las respuestas = B” en un 58,33% y, finalmente, “Identifica las respuestas = R” en un 39,58 %.

[f] ¿Cuál de los siguientes atributos: “Asigna las entradas y salidas a cada proceso = M”, “Define las demoras = M”, “Asigna las entradas y salidas a cada proceso = R”, “Define las demoras = R” incide más sobre el atributo “DFD = M”?

Resultado: El atributo que incide más sobre “DFD=M” es “Asigna las entradas y salidas a cada proceso = M” en un 80%, le sigue “Define las demoras = M” en un 74%, posteriormente “Asigna las entradas y salidas a cada proceso = R” en un 20 %, y finalmente, “Define las demoras = R” en un 18%.

[g] ¿Cuál de los siguientes atributos: “Define las demoras = B”, “Define las demoras = M”, “Asigna las entradas y salidas a cada proceso = R”, “Asigna las entradas y salidas a cada proceso = B”, “Identifica la comunicación entre procesos = M” incide más sobre el atributo “DFD = R”?

Resultado: El atributo que incide más sobre “DFD = R” es “Asigna las entradas y salidas a cada proceso = B” en un 61,11%, luego “Define las demoras = B” en un 55,56%, le sigue “Asigna las entradas y salidas a cada proceso = R” en un 38,89 %, a continuación “Identifica la comunicación entre procesos = M” en un 33,33% y finalmente “Define las demoras = M” en un 11,11 %.

[h] ¿Cuál de los siguientes atributos: “Define las demoras = R”, “Asigna las entradas y salidas a cada proceso = B”, “Identifica la comunicación entre procesos = B”, “Define las demoras = B”, “Identifica la comunicación entre procesos = R”,

“Define las demoras = B” incide más sobre el atributo “DFD = B”?

Resultado: El atributo que incide sobre “DFD = B” es “Define las demoras = R” en un 33,33%, “Asigna las entradas y salidas a cada proceso = B” en un 61,11%, “Identifica la comunicación entre procesos = B” en un 22,22%, “Define las demoras = B” en un 55,56%, “Identifica la comunicación entre procesos = R” en un 44,44 %, “Define las demoras = B” en un 55,56%.

3.2.10. Subpaso 5

En el grupo de reglas que se presentan en la Tabla 1, vinculadas a Diagrama de Contexto (DC), se puede comprobar la fuerte relación que existe entre la variable “Identifica entradas y salidas” y la comprensión del DFD.

En el grupo de reglas que se presentan en la Tabla 2, vinculadas a la Tabla de Eventos, se puede comprobar la fuerte relación que existe entre las variables “Puede Identificar funciones” y la comprensión del DFD.

En el grupo de reglas que se presentan en la Tabla 3, vinculadas al Diagrama de Flujos de Datos (DFD), se pueden realizar las siguientes observaciones:

[i] Se comprueba la gran importancia que tiene la variable “Asigna las entradas y salidas a cada proceso” en la comprensión del DFD. Si bien era un resultado esperado, su expresión en la regla 13 “Si Define las demoras = B y Asigna las entradas y salidas a cada proceso = R”, hizo reflexionar a los docentes de la asignatura ya que se esperaba que el valor fuera DFD = B.

[ii] En función de lo definido en el punto anterior, llama la atención de los docentes de la asignatura que, la regla 14 “Si Define las demoras = R y Asigna las entradas y salidas a cada proceso = R” tenga el resultado DFD = M, cuando se hubiera esperado que asumiera el valor R.

- [iii] La relación de la variable “Identifica la comunicación entre procesos”, con la comprensión del DFD, no es tan fuerte como se hubiera esperado antes de la obtención de las reglas. Esto se visualiza en las reglas 18 y 19.
- [iv] Se comprueba que la variable “Define las demoras”, no impacta en la comprensión del DFD por sí misma, como era de esperar. Se descubre que la relación de la variable con DFD es muy débil, ya que siempre el resultado se ve influenciado por la variable “Asigna las entradas y salidas a cada proceso”. Reglas 12, 14, 15 y 16.

Derivadas de la aplicación del proceso de ponderación de interdependencia de atributos surgen las siguientes observaciones:

- [v] Se verifica la importancia que tiene el atributo *Identificación de Entidades Externas* (93,88%), porque si el alumno lo logra, identifica el contexto del sistema, luego puede identificar entradas y salidas, llegando a definir el Diagrama de Contexto óptimo.
- [vi] No poder *Identificar las respuestas* (79,49%) de una función indica que, el alumno no adquirió la habilidad necesaria para armar la Tabla de Eventos.
- [vii] Si el alumno logra *Identificar las funciones* (97,92%) del sistema indica que apropió bien los conceptos y puede construir una Tabla de Eventos.
- [viii] La *Asignación de entradas y salidas* (61,11%) en cada proceso, propicia que se logre la construcción aceptable del DFD. Que el alumno no logre esta asignación indica que no ha logrado apropiarse de los conceptos necesarios para construir el Diagrama.

6. Conclusiones

En este trabajo se ha presentado un proceso basado en explotación de información, aplicado a la identificación de errores de apropiación de conceptos en estudiantes universitarios.

Se ha presentado una prueba de concepto de la utilización del proceso en el dominio de la asignatura Análisis de Sistemas.

El proceso propuesto constituye una nueva herramienta de diagnóstico en el área; y si bien los resultados obtenidos en este estadio del proyecto son prometedores, no son concluyentes.

El próximo paso de esta investigación será contrastar el proceso en dominios correspondientes a otras asignaturas.

7. Financiamiento

Las investigaciones que se reportan en este artículo han sido financiadas parcialmente por el Proyecto de Investigación 33A105 del Departamento de Desarrollo Productivo y Tecnológico de la Universidad Nacional de Lanús, por el Proyecto de Investigación 40B133 de la Universidad Nacional de Río Negro - Sede Andina (El Bolsón), y por el Proyecto EIUTIBA1121 de la Facultad Regional Buenos Aires de la Universidad Tecnológica Nacional.

8. Referencias

- Aebli, H. (1991) *Factores de la enseñanza que favorecen el aprendizaje autónomo*. Narcea S.A.
- Britos, P. (2005) *Objetivos de Negocio y Procesos de Minería de Datos Basados en Sistemas Inteligentes*. Reportes Técnicos en Ingeniería del Software.. 7(1): p. 26-29.
- Britos, P. (2008). *Procesos de Explotación de Información Basados en Sistemas Inteligentes*. Tesis de Doctorado en Ciencias Informáticas. Facultad de Informática. Universidad Nacional de La Plata.
- Britos, P., García-Martínez, R. (2009). *Propuesta de Procesos de Explotación de Información*. Proceedings XV Congreso Argentino de Ciencias de la Computación. Workshop de Base de Datos y Minería de Datos. Págs. 1041-1050. ISBN 978-897-24068-4-1.

- Britos, P., Jiménez Rey, E., García-Martínez, E. (2008). *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.
- Britos, P., Hossian, A., García Martínez, R., Sierra, E.(2005) *Minería de Datos Basada en Sistemas Inteligentes*. Nueva Librería.
- Fayyad U.M., Piatetsky Shapiro G., Smyth P. 1996 *From Data Mining to Knowledge Discovery: An Overview*. Advances in Knowledge Discovery and Data Mining. AAAI/MIT Press, 1996. p 1-34.
- 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 Congreso Argentino de Ciencias de la Computación, Workshop de Tecnología Informática Aplicada en Educación, Artículo 1881. ISBN 978-987-24611-0-2.
- Jiménez Rey, E., Rodríguez, D., Britos, P., García-Martínez, R. (2009). *Caracterización de Problemas de Aprendizaje Basada en Explotación de Información*. Proceedings XI Workshop de Investigadores en Ciencias de la Computación. Pág. 627-629. ISBN 978-950-605-570-7.
- Jiménez Rey, E., Rodríguez, D., García-Martínez, R. (2010). *Mapa de Aprendizajes Significativos como Modelo de Representación de Conocimientos Previos en el Proceso de Caracterización de Problemas de Aprendizaje*. Proceedings del V Congreso de Tecnología en Educación y Educación en Tecnología. Pág. 547-555. ISBN 978-987-1242-42-9.
- Larose, D. T. (2005). *Discovering Knowledge in Data, an introduction to Data Mining*. John Wiley & Sons
- Spotts, T y Bowman, M. (1995). *Faculty Use of Instructional Technologies in Higher Education*, en *Educational Technology*, 35(2): 56-64.