

Descubrimiento de Patrones cognitivos en Evaluaciones de Informática basado en Explotación de Información

Andrea Laluf, Leandro Saavedra, Laura C. Díaz, Carlos A. Bartó, Aldo M. Algorry
Departamento Computación, Facultad de Ciencias Exactas, Físicas y Naturales, Universidad Nacional de Córdoba
andrealalufgonzalez@gmail.com, leandro.saavedra19@gmail.com, lcd_ic@yahoo.com.ar,
cbarto@gmail.com, aalgorry@gmail.com

Resumen

Este trabajo es parte de un proyecto abocado a la aplicación de técnicas de Inteligencia Artificial en los procesos de evaluación a estudiantes de Programación de las carreras de Ingeniería. Se muestran los primeros resultados de la aplicación de aprendizaje inductivo para encontrar relaciones entre las respuestas dadas por estudiantes de Informática en sus diferentes instancias de evaluación para acreditar la asignatura. El objetivo consiste en indagar la existencia de relaciones en los procesos cognitivos, materializados en las respuestas dadas por los estudiantes en la secuencia de Evaluaciones durante la cursada. Se espera que estos resultados contribuyan a la mejora de los procesos de aprendizaje y evaluación a raíz de la oportunidad de realizar predicciones del comportamiento académico futuro del individuo en proceso de aprendizaje, basado en el descubrimiento de estos patrones de respuesta y en base a esto poder determinar de manera personalizada las necesidades de refuerzo en los contenidos.

Palabras clave: Patrones de Respuesta, Aprendizaje de Programación, Descubrimiento de Información basado en Datos.

1. Introducción: Contexto y Presentación del problema

1.1 Contexto

El presente fue realizado mayormente por dos estudiantes de grado de la carrera de Ing. en Computación con el soporte de los

investigadores del equipo, por lo que uno de los objetivos subyacentes de este trabajo es la formación de futuros investigadores a través de la práctica científica. Fue elaborado en el marco de la actividad de educación continua para egresados con modalidad taller, titulado: 'Modelos de Evaluación en Informática', organizado por el Departamento Computación de esta Unidad Académica, desarrollado durante el segundo cuatrimestre del año 2014, dirigido a docentes e investigadores en Educación en Programación. Esta actividad, a su vez, es una de las acciones previstas en el proyecto de investigación acreditado por SECyT de la Universidad Nacional de Córdoba para el bienio 2014-2015: 'Inteligencia Computacional y TIC: Estrategias para Facilitar el Aprendizaje en la Universidad Nacional de Córdoba'. Todos los autores de este trabajo, estudiantes y docentes, forman parte del equipo de investigación.

1.2 Presentación del Problema. Importancia

El uso Tecnologías Inteligentes de Explotación de Información ofrece la oportunidad de descubrir relaciones de tipo socioeconómicas, académicas y cognitivas, de los individuos en proceso de aprendizaje, que con otras metodologías no serían necesariamente detectadas [Kuna, 2010]. Estos hallazgos son de utilidad para el diseño de políticas públicas, alimentar las bases de conocimiento de Sistemas Tutores Inteligentes, tomar acciones preventivas y correctivas por parte de los actores involucrados en forma directa en los procesos de aprendizaje y evaluación, etc.; en contextos de masividad [Díaz, 2014] Además, la detección temprana de las capacidades de los estudiantes resultó un

factor significativo en la mejora del aprendizaje de Programación, estos hallazgos realizados en [Bornat et al, 2008], despiertan el interés en el uso de Sistemas de Redes Neuronales Artificiales y otras Tecnologías Inteligentes para predecir su rendimiento académico en las carreras de Ingeniería.

En esta oportunidad, el problema se centra en una de las instancias de acreditación de conocimientos de la asignatura Informática, en el primer año de las carreras de Ingeniería en esta Facultad.

Las denominadas Evaluaciones Conceptuales, siete en total hasta 2014, tomadas a los estudiantes bajo la modalidad de evaluación continua sumativa, fueron objeto de tratamiento para incorporar mejoras, durante el periodo 2012-2013, en el proyecto de investigación que precede a éste: ‘Sistemas Inteligentes Aplicados a la Enseñanza de la Programación en Ingeniería’ [Bartó et al, 2013]. Se introdujeron mejoras en el sentido de diseñar, desarrollar, implementar y medir un Sistema de Evaluaciones con preguntas (ítems) con independencia local, bajo la modalidad de opción múltiple dicotómicas, con refuerzo. Para el desarrollo de cada una de las opciones de respuesta, consideradas como modelos de respuesta errónea del estudiante (MRE), se interpretó por expertos del dominio y se desarrolló el algoritmo codificado que expresara el comportamiento cognitivo de cada uno de los MRE para cada ítem del total que componen el banco de preguntas de la asignatura [Díaz et al, 2013].

A partir de estas respuestas, el objetivo de esta presentación es encontrar relaciones entre las respuestas dadas por los estudiantes en sus instancias evaluativas, como una primera acción hacia la predicción de comportamiento cognitivo de los estudiantes basado en sus conocimientos previos, lo cual facilitará la implementación de acciones para la mejora en el aprendizaje con refuerzo personalizado en contextos de masividad.

1.1 Objetivos de la investigación

Mediante el uso de algoritmos inteligentes de aprendizaje inductivo, se persigue el objetivo

de encontrar relaciones entre las respuestas a los ítems de una evaluación conceptual (EC) con las respuestas a un ítem de otra EC, tomada con posterioridad. En otras palabras, obtener una primera aproximación al descubrimiento de patrones cognitivos en las respuestas dadas por los individuos sujetos a evaluación, en Informática, para las carreras de Ingeniería de esta Unidad Académica, que utilizan a Python como lenguaje de programación.

2. Desarrollo

En la sección 2.1 se realiza una breve descripción de la modalidad de Evaluaciones Conceptuales en Informática, con las mejoras introducidas, mencionadas en el apartado anterior. En el apartado 2.2 se realiza una breve reseña del modo de funcionamiento de los algoritmos de aprendizaje inductivo TDIDT. En el apartado 2.3 se describen los aspectos metodológicos.

2.1 Descripción del Modelo de evaluación

En la plataforma MOODLE, para los estudiantes que cursan Informática correspondientes a las carreras de Ingeniería Civil, Mecánica, Mecánica Electricista, Industrial, Aeronáutica y Química, se estructuraron las siete evaluaciones, objeto de análisis. Contienen nueve ítems cada una, distribuidos a lo largo del libro de texto de la asignatura para el aprendizaje en lenguaje Python [Marzal, A y García, I, 2003].

Las preguntas –ítems- se diseñaron con cinco opciones de respuestas, cuatro de ellas correspondían a los modelos de respuesta erróneos detectados en las evaluaciones tomadas durante 2012. Las preguntas retroalimentan las respuestas conforme a los MRE, con lo cual se pretende dar al estudiante una explicación de sus conceptos erróneos, facilitando así la disminución de la brecha entre el modelo conceptual y su interpretación. Por razones de seguridad se generaron alrededor de 30 preguntas promedio, conceptualmente equivalentes, con diferentes valores de entrada y salida, para enfrentar a los estudiantes a evaluaciones del mismo tipo y

nivel de complejidad, con idénticos MRE pero con valores diferentes.

2.1 Árboles de inducción

El aprendizaje inductivo es un caso particular entre las técnicas de aprendizaje a partir de ejemplos, siendo su principal objetivo el inducir reglas a partir de los datos disponibles, para lo cual procede a clasificar en la clase correspondiente diferentes objetos, basándose en el valor de las características o atributos que los definen.

Para la aplicación de estos sistemas de aprendizaje inductivo se parte de un conjunto de ejemplos donde cada ejemplo debe tener la misma estructura, consistente en una conclusión (o decisión) y un número de características o atributos que definen esa conclusión o decisión.

El sistema construye un árbol de decisión que representa la relación existente entre la conclusión-decisión y sus atributos. Es decir, se produce un proceso de generalización de forma que el árbol de decisión generado clasifica correctamente los ejemplos dados. Este árbol, además, se caracteriza por ser el óptimo en el sentido que minimiza el número de atributos requeridos para alcanzar la conclusión-decisión, siendo esta la explicación de por qué ciertos atributos no aparecen en el árbol.

En terminología de árboles de decisión, la base del árbol se denomina Raíz. Cada división para un atributo representa una Rama y el punto final del árbol donde se alcanza una decisión es denominado Hoja.

Estos sistemas de aprendizaje inductivo se basan, generalmente, en los algoritmos ID3 y C4.5. El ID3 es un algoritmo simple pero sin embargo, potente, cuya finalidad es la elaboración de un árbol de decisión bajo las siguientes premisas:

1. Cada nodo corresponde a un atributo y cada rama al valor posible de ese atributo. Una hoja del árbol especifica el valor esperado de

la decisión de acuerdo con los ejemplos dados. La explicación de una determinada decisión viene dada por la trayectoria desde la raíz a la hoja representativa de esa decisión.

2. A cada nodo es asociado aquel atributo más informativo que aún no haya sido considerado en la trayectoria desde la raíz.

3. Para medir cuan informativo es un atributo se emplea el concepto de entropía: cuanto menor sea el valor de la entropía, menor será la incertidumbre y más útil será el atributo para la clasificación.

El ID3 es capaz de tratar con atributos cuyos valores sean discretos o continuos. En el primer caso, el árbol de decisión generado tendrá tantas ramas como valores posibles tome el atributo. Si los valores del atributo son continuos, el ID3 no clasifica correctamente los ejemplos dados. Por ello, [Quinlan, 1993] propuso el C4.5, como extensión del ID3, que permite:

1. Construir árboles de decisión cuando algunos de los ejemplos presentan valores desconocidos para algunos de los atributos.

2. Trabajar con atributos que presenten valores continuos.

3. El árbol de decisión ha sido construido a partir de un conjunto de ejemplos, por tanto, reflejará correctamente todo el grupo de casos. Sin embargo, como esos ejemplos pueden ser muy diferentes entre sí, el árbol resultante puede llegar a ser bastante complejo, con trayectorias largas y muy desiguales. Para facilitar la comprensión del árbol puede realizarse una poda del mismo, lo que significa la sustitución de una parte del árbol (sub-árbol) por una hoja. La poda tendrá lugar si el valor esperado de error en el sub-árbol es mayor que con la hoja que lo sustituya.

2.3 Aspectos Metodológicos

A partir de la información disponible en la base de datos de Moodle, que contiene información de distintos cursos de varias carreras que se dictan en esta facultad, fue necesario realizar un filtrado para acceder a los cursos de Informática, lenguaje Python, ambos

semestres de 2013 y 2014. Esta consulta se realizó en MySQL. El resultado fue una tabla con un registro por estudiante, que contiene campos que indican la respuesta a cada uno de los nueve ítems de las siete EC que conforman el modelo de evaluación descrito en el apartado 2.1.

Se escogieron todas los ítems de la EC 3 para ser relacionadas con una de las preguntas que componen la EC 5, esta decisión fue tomada atendiendo a las relaciones de conceptos que guardan entre sí ambas evaluaciones, a criterio de los expertos, docentes e investigadores de este equipo. Así, se usaron las nueve respuestas dadas por cada alumno en la EC3, como atributos de entrada y la respuesta dada en el ítem escogido de la EC5, como atributo clase. La distancia en la aplicación de las dos evaluaciones conceptuales durante el desarrollo de las clases, es usualmente de tres semanas.

Es necesario destacar que, la variabilidad en las instancias de las semillas usadas en la construcción de las preguntas introduce un ruido que dificulta la confiabilidad de los resultados.

Se transcribe una de las diez preguntas de la EC3 que se utilizaron como dataset:

P1: 4.2 Sentencias iterativas La sentencia while 1a (cbT)

Indicar el resultado de la ejecución del siguiente programa:

```
u=7
v=7
s=1
t=4
while v<t
    v+=s
    u=v
print u
```

respuestas posibles

- 7 respuesta correcta
 4 Evalúa mal la condición del `while`, no toma en cuenta la actualización de `v`
 3 Evalúa mal la condición del `while`, y considera al `+=` y al `-=` como asignaciones simples
 -3 Evalúa mal la condición del `while`, y lo toma como un `if`
 10 Evalúa mal la condición del `while`, lo toma como un `if` y no se aplica el operador `-=` para actualizar `u`

Ilustración 1. Imagen de la Pregunta 1 de la EC3, como se presenta al estudiante en el aula virtual de Informática bajo la plataforma Moodle

3. Resultados

Si bien incipientes, se muestran los resultados obtenidos para el tratamiento de la base de datos como modelo de respuesta dicotómica (sección 3.1) y, además, algunos de los resultados para cada MRE de la pregunta clase correspondiente a la EC5 (sección 3.2).

3.1 Resultados para el modelo dicotómico

Se muestra un reporte de la aplicación Tanagra:

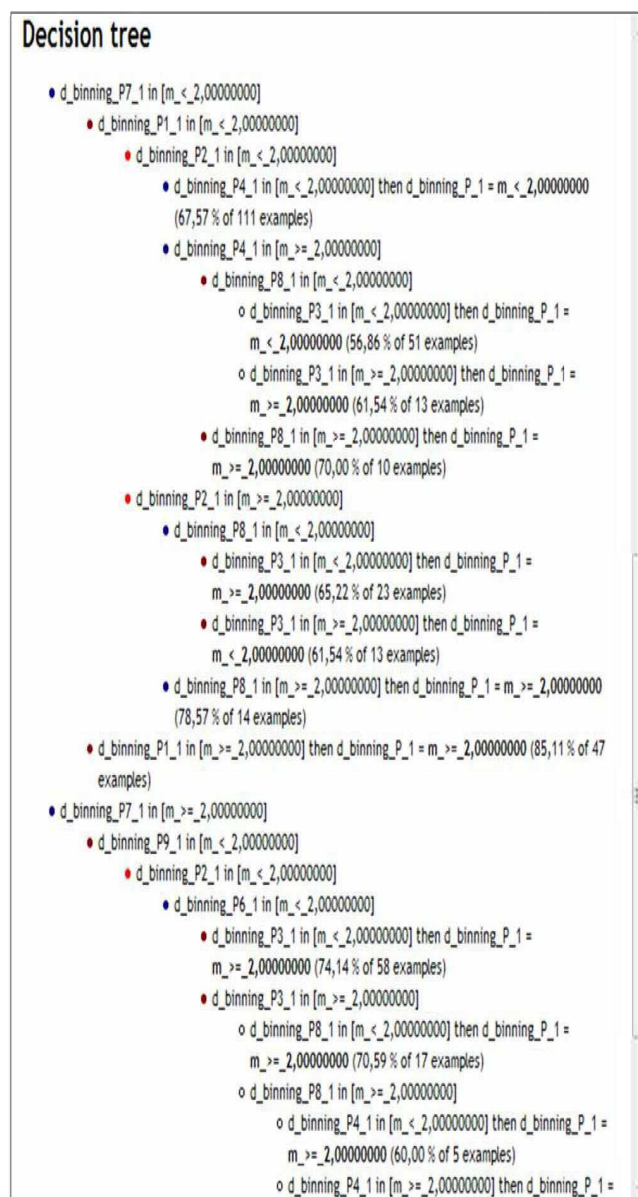


Ilustración 2. Reporte en Tanagra para el modelo dicotómico de respuestas.

Las reglas obtenidas con un error del 24,28% utilizando el algoritmo C4.5 con los atributos en números binarios son las siguientes:

- El 70% de diez ejemplos de los alumnos que respondieron bien las preguntas 7 ,1 y 2 y mal las preguntas 4 y 8 de la EC3, respondieron mal la pregunta de la EC5.
- El 78,57% de catorce ejemplos de los alumnos que respondieron bien a las preguntas 7 y 1 y respondieron mal a las preguntas 2 y 8 de la EC3, respondieron mal la pregunta de la EC5.
- El 85,11% de cuarenta y siete ejemplos de los alumnos que respondieron bien a la pregunta 7 y mal a la pregunta 1 de la EC3, respondieron mal a la pregunta de la EC5.
- El 74,14% de cincuenta y ocho ejemplos de los alumnos que respondieron mal a la pregunta 7 y bien a las preguntas 9, 2, 6 y 3 de la EC3 respondieron mal a la pregunta de la EC5.
- El 70,59% de diecisiete ejemplos de los alumnos que respondieron mal a las preguntas 7 y 3 y respondieron bien a las preguntas 9, 2, 6 y 8 de la EC3, respondieron mal a la pregunta de la EC5.
- El 71,43% de siete ejemplos de los alumnos que respondieron mal a las preguntas 7, 3, 8 y 4 y respondieron bien a las preguntas 9, 2 y 6 de la EC3, respondieron bien a la pregunta de la EC5.
- El 88,24% de diecisiete ejemplos de los alumnos que respondieron mal a las preguntas 7 y 6 y bien la las preguntas 9 y 2 de la EC3 respondieron mal a la pregunta de la EC5.
- El 83,33% de noventa ejemplos de los alumnos que respondieron mal la pregunta 7 y 2 y bien la pregunta 9 de la EC3 respondieron mal la pregunta de la EC5.
- El 88,50% de ciento trece ejemplos de los alumnos que respondieron mal las preguntas 7 y 9 de la EC3 respondieron mal a la pregunta de la EC5.

Estos resultados estarían indicando que los alumnos que respondieron mal a la pregunta 7, una de las preguntas que trabaja con el ciclo for-in, entonces también respondieron mal la

pregunta de la EC5. El caso en el que respondieron bien la pregunta habiendo respondido mal la pregunta 7 contiene muy pocos ejemplos.

3.2 Algunos resultados de los MRE relevantes



Ilustración 3. Ejemplo de Reporte en Tanagra con todos los MRE del estudiante.

Se utilizó el algoritmo C4.5 con un error del 21,69% con todas las preguntas de la EC3 con sus respuestas correspondientes y como target la pregunta de la EC5 con cada uno de los modelos de respuesta:

El 95,24% de 21 ejemplos de los alumnos que respondieron a las preguntas de la evaluación EC3 de la siguiente manera:

- No respondió a la pregunta 7 con la opción correcta
- No respondió a la pregunta 9 con el MRE 4 que dice que se está mostrando una variable que no es la pedida y evaluando mal la condición.

- No respondió a la pregunta 6 con el MRE 3 que dice que comenzó el ciclo con un número superior al que debería.
- Respondió a la pregunta 4 con el MRE 4 que dice lo siguiente: Considera que los rangos dos los bucles for inician en 1 y terminan en $m+1$ y $n+1$.

Entonces no respondió a la pregunta de la EC5 con la opción correcta.

El 95,24% de 21 ejemplos de los alumnos que respondieron a las preguntas de la evaluación EC3 de la siguiente manera:

- No respondió a la pregunta 7 con la opción correcta
- No respondió a la pregunta 9 con el MRE 4 que dice que se está mostrando una variable que no es la pedida y evaluando mal la condición.
- No respondió a la pregunta 6 con el MRE 3 que dice que comenzó el ciclo con un número superior al que debería.
- Respondió a la pregunta 4 con el MRE 4 que dice lo siguiente: Considera que los rangos dos los bucles for inician en 1 y terminan en $m+1$ y $n+1$.

Entonces no respondió a la pregunta de la EC5 con la opción correcta.

El 100% de 13 ejemplos de los alumnos que respondieron a las preguntas de la evaluación EC3 de la siguiente manera:

- No respondió a la pregunta 7 con la opción correcta
- Respondió a la pregunta 9 con el MRE 4 que dice que se está mostrando una variable que no es la pedida y evaluando mal la condición.

Entonces no respondió a la pregunta de la EC5 con la opción correcta.

A continuación, algunas reglas obtenidas de la aplicación del dataset donde se tienen todas las preguntas con cada modelo de respuesta de la EC3 y como objetivo a uno de las opciones de respuesta de la pregunta de la EC5.

El 73,00% de 100 ejemplos de los alumnos que respondieron a las preguntas de la evaluación EC3 de la siguiente manera:

- Respondió a las preguntas 7 y 1 con la opción correcta.
- No respondió a la pregunta 5 con la opción correcta.
- No respondió a la pregunta 3 con el MRE 5 que dice: Evalúa incorrectamente ambas comparaciones.
- No respondió la pregunta 8 con el MRE 2 que dice: Muestra el valor de la variable actualizada y no el resultado.
- No respondió a la pregunta 4 con el MRE 5 que dice: Considera que el valor inicial de la variable declarada es 0.
- No respondió a la pregunta 4 con el MRE 4 que dice: Considera que los rangos dos los bucles for inician en 1 y terminan en $m+1$ y $n+1$.
- Respondió a las preguntas 9 y 3 con la opción correcta.
- No respondió a la pregunta 2 con el MRE 2 que dice: La condición del while se ha tomado como $==$.
- No respondió a la pregunta 3 con el MRE 4 que dice: Ignora todas las estructuras de decisión.
- No respondió a la pregunta 2 con el MRE 5 que dice: Se considera que sin importar el caso el resultado es siempre 'd'.
- No respondió a la pregunta 5 con el MRE 5 que dice: Ejecutó una vez el ciclo for, para $i=0$, y en el ciclo while hizo una iteración menos.

Entonces respondió a la pregunta de la EC5 con la opción correcta.

Se utilizó el algoritmo C4.5 con un error del 14,58%. Para el MRE 2 de la EC5, que dice 'Los for se encuentran anidados', el 100.00% de 8 ejemplos de los alumnos que respondieron a las preguntas de la evaluación EC3 de la siguiente manera:

- Respondió a la pregunta 6 con el MRE 2 que dice: 'Realizó una iteración demás.'

Entonces no respondió a la pregunta de la EC5 con el MRE 2.

Se utilizó el algoritmo C4.5 con un error del 8,64% con todas las preguntas de la EC3 con sus modelos de respuestas correspondientes y como atributo clase a la pregunta de la EC5 con el MRE 5 que dice: Recorrer una lista no produce sublistas sino los valor de los elementos de la lista, en este caso son números por lo que deben sumarse no concatenarse.

El 100.00% de 32 ejemplos de los alumnos que respondieron a las preguntas de la evaluación EC3 de la siguiente manera:

- No respondió la pregunta 4 con el MRE 4 que dice: Considera que los rangos dos los bucles for inician en 1 y terminan en m+1 y n+1.
- No respondió la pregunta 6 con el MRE 2 que dice: Realizó una iteración demás, incluyó a 10.
- Respondió la pregunta 3 con el modelo mental 5 que dice lo siguiente: Evalúa incorrectamente ambas comparaciones.

Entonces no respondió la pregunta de la EC5 con el MRE 5.

Se utilizó el algoritmo C4.5 con un error del 3% con todas las preguntas de la EC3 con sus MR correspondientes y como target la pregunta de la EC5 con el MRE 3 que dice que ‘Se suman los valores de los elementos no sus posiciones’:

- Este modelo no fue utilizado por 22 alumnos únicamente de los 590 que es el total y la regla obtenida es que el 96,27% de 590 ejemplos de los alumnos que no respondieron la pregunta 3 con el MRE 1 que es el correcto, entonces no respondieron a la pregunta de la EC5 con el MRE 3.

Conclusiones

De los resultados obtenidos, se deduce que el hallazgo más importante es una primera confirmación de la hipótesis que existe una correlación cuantificable entre los resultados que obtiene un alumno en una evaluación y los resultados que obtiene en una evaluación posterior con dependencia entre ambos

contenidos evaluados. Esto se presenta como patrones cognitivos que evidenciados en las reglas obtenidas a través de los árboles de inducción.

Esto permitiría realizar estudios a nivel cognitivo para predecir el comportamiento o bien para mejorar el pronóstico mediante acciones personalizadas o implementar técnicas de aprendizaje automático para predecir deficiencias futuras en las evaluaciones de los exámenes.

Futuros Pasos

A partir de estos resultados se pueden profundizar los estudios para determinar en qué medida la correlación temática y las características personales del estudiante son significativas para predecir el rendimiento académico.

Si bien son los primeros resultados, se hallaron pocas reglas significativas y la base de datos es pequeña para el nivel de ramificación y profundidad del árbol como para obtener ejemplos en número relevante de entradas por clase se abre una interesante línea de acciones a futuro. Se podría, a partir de estos resultados, construirse reglas con incertidumbre para conformar un sistema experto.

Otras propuestas sugieren trabajar con redes neuronales de tipo Perceptrón multicapa con aprendizaje supervisado por corrección del error, capaces de aprender para generalizar y predecir.

Además, también se cuenta con los datos de 2014 que podrían usarse para validar los resultados obtenidos.

De cualquier modo, lograr, aunque escasos, resultados que permitan producir patrones de comportamiento cognitivo, proporcionando los modos de respuesta de los estudiantes, materializadas en estos primeros resultados, es un hallazgo significativo que invita a generar acciones útiles a partir de ellas y a continuar el análisis.

Bibliografía

Bartó, C., Díaz, L. 2013. *Intelligent Systems Applied to Computer Engineering Teaching*. IEEE Latin America Transactions, 11(1): 591-595. ISSN 1548-0992.

Bornat, R., Dehnadi S. y Hamilton, S. 2008. *Mental models, Consistency, and Programming Aptitude*. Australian Computer Society. ACE.

Díaz, L., Algorry, A., Eschoyez, M., Barto, C., Marangunic, R. 2013. *Actions towards the application of intelligent systems in computer education*. IEEE Latin America Transactions, 11(1): 591-595. ISSN 1548-0992.

Díaz, L. 2014. *Investigación en Progreso: Gestión de la Educación Superior en Contextos de Masividad Basada en Tecnologías Inteligentes de Transformación de*

Información. Revista Latinoamericana de Ingeniería de Software, 2(1): 53-76. ISSN 2314-2642.

Marzal, A y García, I, 2003. *Introducción a la Programación con Python*. Departamento de Lenguajes y Sistemas Informáticos. Universitat Jaume I.

Kuna, H., García Martínez, R. Villatoro, F. 2010. *Pattern Discovery in University Students Desertion Based on Data Mining*. Advances and Applications in Statistical Sc. J., 2(2): 275-286. ISSN 0974-6811.

Quinlan, J. R. 1993. *C4.5: Programs for Machine Learning*. San Mateo: Morgan Kaufmann.