

HACIA LA DETECCIÓN DE LOS MODELOS MENTALES DE LOS ESTUDIANTES DE PROGRAMACIÓN

Laura Cecilia Díaz, Rodolfo Marangunic, Carlos Alberto Bartó

lcd_ic@yahoo.com.ar, rmarangunic@efn.uncor.edu, cbarto@gmail.com

Laboratorio de Educación Virtual, Departamento de Computación

F. C. E. F. y N. - U. N. C.

Resumen

En el marco del proyecto acreditado por SECYT que lleva adelante el Laboratorio de Educación Virtual del Departamento de Computación: 'Sistemas inteligentes aplicados a la enseñanza de la programación en Ingeniería', se presentan los primeros resultados referidos a la detección de los modelos mentales de los estudiantes conducentes a proveer la base de conocimiento a un Tutor Virtual Inteligente, el cuál vía retroalimentación le facilitará al estudiante la construcción de modelos conceptuales.

Se describe la metodología empleada para la detección de los modelos mentales de los alumnos que cursaron la asignatura Informática durante el primer cuatrimestre 2012. Esta información alimentará la base de conocimiento del Sistema Tutor Inteligente SIETTE, que la cátedra pretende utilizar para mejorar la retroalimentación de su aula virtual y el sistema de evaluación automático actual. Se muestran el tratamiento estadístico, sus resultados e interpretaciones.

Por otra parte, se muestran nuevos resultados de la aplicación de Redes Neuronales Artificiales en la detección temprana de las capacidades del estudiante para enfrentar con éxito el aprendizaje de la programación. Estos se corresponden con los primeros cursos de acción del equipo, antecedentes del proyecto. En tal sentido, es interesante observar la serie de tiempo a partir del 2010.

La contribución más importante de esta presentación es el análisis confirmatorio, mediante test de hipótesis, del grupo testigo de preguntas –ítems- que se utilizó como

información de modelos mentales para construir la base de conocimiento de SIETTE. De este análisis se desprenden las conclusiones que sugieren modificaciones sobre los distintos cursos de acción, tanto para el proyecto de investigación como para el sistema de evaluación que se utiliza actualmente para la acreditación de la asignatura. Se están llevando a cabo estas acciones correctivas que deben ponerse a punto antes del inicio del próximo ciclo lectivo.

Palabras clave: Tutor Inteligente, Modelos Mentales, Conocimiento, Predicción, Rendimiento.

Introducción

Este trabajo es parte de un proyecto abocado a incorporar Sistemas Tutores Inteligentes en la enseñanza de la Programación en Ingeniería como estrategia para mejorar el rendimiento académico de los estudiantes de Ingeniería.

Una de las líneas de acción que se llevó a cabo durante el primer año fue la investigación exploratoria, tanto de los STI que han sido utilizados en experiencias similares, como de los modelos y estrategias de aprendizaje que mejor se adaptan a cursos de Programación, ello sin descuidar la necesaria compatibilidad entre ambos a la hora de pensar en su implementación.

En tal sentido, los criterios que prevalecieron para la elección de los STI, además de la accesibilidad, se orientan a la mejora de la calidad de los procesos de aprendizaje y de evaluación. Para el primero se pensó en acciones de retroalimentación a partir de una base de conocimiento más alineada con las dificultades reales del estudiante. Para mejorar el proceso de evaluación, las exigencias se orientaron hacia dos aspectos relevantes: el aprovechamiento del aprendizaje durante esta instancia y a obtener resultados que mejor expresen el nivel de conocimiento adquirido por el estudiante. Se escogió a SIETTE (Bartó, Díaz: 2012) como STI que satisface adecuadamente ambos requerimientos.

Por otra parte, y desde un punto de vista estrictamente cognitivo, nuestro equipo de trabajo ha adherido a un modelo (Johnson-Laird:1993), para lo cual se hace

necesario pensar en metodologías que permitan detectar los modelos mentales de los estudiantes (MME). Estos modelos mentales nutren la base de conocimiento de SIETTE para mejorar el proceso de evaluación.

Es a través de la estrategia de retroalimentación que se busca achicar la brecha entre los MME y el modelo conceptual del conocimiento objeto de aprendizaje. Trabajando en esta dirección, se hace necesario realizar experimentos sobre grupos de estudiantes con el fin de detectar los MME y reconstruir un proceso de evaluación que, atendiendo a estos, posibilite la retroalimentación para favorecer el aprendizaje durante la instancia de evaluación y además, que refleje mejor el nivel de conocimiento adquirido.

El Proceso de acreditación de la asignatura consta actualmente de tres partes: las Evaluaciones Conceptuales, la Especificación de Algoritmo y los Trabajos Prácticos. Las dos últimas no son actualmente objeto de incorporación a SIETTE. Sin embargo, se están llevando a cabo otro tipo de acciones, cuyos relatos son objeto de presentación en este evento.

Las Evaluaciones Conceptuales, adaptables a SIETTE, están organizadas en veinte ítems distribuidos en siete test que se aplican a lo largo del desarrollo de la asignatura, con lo cual es posible la retroalimentación para el aprendizaje. Los contenidos abarcan todos los aspectos conceptuales desarrollados. El estudiante debe alcanzar en esta epata un mínimo de sesenta por ciento como condición excluyente para la acreditación.

Se describen los aspectos metodológicos relacionados con la provisión de la base de conocimiento para habilitar el proceso de Evaluaciones Conceptuales con SIETTE en 2013. Esto incluye el análisis confirmatorio mediante test de hipótesis, del grupo testigo de preguntas –ítems- que se utilizó como información de modelos mentales para construir dicha base de conocimiento, en adelante Grupo B. Además se muestran el análisis y la interpretación de resultados a partir de los cuales se sugieren modificaciones, tanto para las acciones de investigación como para el sistema de evaluación que se utiliza actualmente en la asignatura.

Por último, la detección temprana de patrones de comportamiento del rendimiento académico es otra contribución a los modelos de aprendizaje de los

estudiantes para la aplicación de STI. En este sentido, el objetivo es predecir rendimiento académico utilizando una red neuronal de aprendizaje supervisado cuya información de entrada es la correspondiente a los test predictores tomados al inicio de cada cursada. Se pretende investigar si es factible proponer la existencia de patrones de comportamiento y que, además éstos sean detectados a través de esta herramienta de la inteligencia computacional.

Desarrollo

En la asignatura Informática se cuenta con un generador de preguntas y un sistema de corrección automático, todo adecuado a la plataforma Moodle. Para cada ítem de cada EC se cuenta con un banco de preguntas con distribución random para garantizar la individualidad de la prueba que le corresponde a cada alumno. Las respuestas son abiertas.

La base de datos de respuestas de los alumnos de Ingeniería que cursaron la asignatura durante los años 2010, 2011 y 2012 contiene la información necesaria para detectar los MME.

En ese sentido, se utilizaron los resultados de las evaluaciones conceptuales de los estudiantes de Informática durante el cursado correspondiente al primer cuatrimestre 2012. Se eligió un grupo de individuos para formar el Grupo B, de 136 alumnos que equivale al 25% de 546 alumnos distribuidos en ocho comisiones, ocho docentes, tres bandas horarias de dictado, todas las carreras de Ingeniería que cursan la asignatura durante el primer cuatrimestre con el lenguaje Python como herramienta de aprendizaje. El 75% de alumnos restantes constituyó lo que en adelante se denomina Grupo A.

Al Grupo B se le aplicó un único instrumento evaluador para cada EC y cada ítem, extraído del banco de preguntas mencionado. Al finalizar la cursada se contaba con los resultados de todas las instancias de evaluación para ambos grupos.

Con esos resultados se llevaron a cabo dos líneas de acción necesarias: la detección de los modelos mentales de los estudiantes y el análisis confirmatorio mediante test de hipótesis entre las preguntas seleccionadas para el Grupo B y el Grupo A con el fin de validar los MME detectados para los conceptos involucrados.

La primera tarea fue realizada por expertos, docentes e investigadores de nuestro equipo de trabajo, e implicó un gran esfuerzo para comprender los procesos de argumentación que conducían a dar esas respuestas. Se propuso detectar para cada una, no más de cuatro MME.

En el procesamiento estadístico, atendiendo al tipo de variable utilizado en el calificador de Moodle, se realizó un test de hipótesis de diferencia entre proporciones y una prueba Chi Cuadrada como respaldatoria del primero.

La siguiente etapa consistirá en aplicar evaluaciones de opción múltiple cuyas respuestas se adecuen a los MME detectados. De este modo se lograrán dos objetivos fundamentales consecuentes con la implementación exitosa de STI: la retroalimentación del estudiante durante los procesos de evaluación y la provisión de datos para detectar automáticamente, mediante Redes Neuronales Artificiales, patrones asociados a los MME y al rendimiento académico.

En relación con este último objetivo, los avances realizados hasta el momento se refieren a la predicción del rendimiento académico de los estudiantes, a partir de los test predictores tomados (Carretero y otros: 2005) y de las calificaciones obtenidas al finalizar el curso.

En esta oportunidad se muestran resultados para el primer cuatrimestre 2012 utilizando una RNA de arquitectura multicapa, tipo Perceptron con aprendizaje supervisado con propagación hacia atrás por corrección del error, de la que ya se cuenta con resultados que fueron motivo de publicaciones anteriores (Díaz y otros:2012) con información correspondiente a los años 2010 y 2011.

Sin embargo, creemos que el verdadero avance en este sentido será la implementación de una RNA con aprendizaje no supervisado del tipo Kohonen que permitirá la detección de patrones de aprendizaje y además relacionarlos con el rendimiento académico de los estudiantes. Se espera contar con avances en esta dirección hacia fines de 2013.

Resultados

El proceso de detección de los Modelos Mentales de los Estudiantes se llevó a cabo mediante el análisis de las respuestas abiertas que los alumnos del Grupo B

daban a cada ítem. Se buscaron respuestas repetidas y además que el modelo mental fuese comprensible, para ello se hicieron importantes esfuerzos en pos de detectar el proceso mental que conducía a la respuesta errónea.

De las veinte preguntas, en trece fue posible interpretar al menos cuatro MME. Sin embargo, en situaciones como las de los ítems 7, 13, 15, 17 y 20, para las cuales no resultó posible entender la semántica de las respuestas erróneas de los estudiantes, surgió la necesidad de recrearlos, apelando para ello a la experiencia que los docentes han adquirido en las observaciones realizadas a los alumnos durante el desarrollo de las clases y en las evaluaciones parciales y finales.

También se detectaron respuestas que, aunque repetidas, resultó imposible reconstruir su semántica como por ejemplo en los ítems 16 y 19.

Por último, se redactó para cada ítem y cada MME, una explicación conducente a destacar el proceso de argumentación errónea detectado. Esto para alimentar la base de conocimiento necesaria en la retroalimentación de las evaluaciones automáticas.

Para utilizar los MME hallados es necesario realizar previamente un análisis confirmatorio, para ello se propuso un test de hipótesis de diferencia entre proporciones de los resultados encontrados correspondientes a los Grupos A y B.

En la figura se muestra un resumen del procesamiento estadístico realizado. La primera columna corresponde a cada una de las siete instancias de evaluación, la segunda indica el número de ítem o

Test estadístico: rechazo
Estadísticamente bueno
Aceptable, puede rechazarse

RESULTADOS EVALUACIONES CONCEPTUALES 1-2012							
EC	Item	Gr B p=X/n	Gr Ct p=X/n	Dif entre proporc Interv conf 99%		X cuadrada	Decisiones
1	1	0,70	0,60	-0,002	0,202	5,74	ok
1	2	0,43	0,56	0,026	0,243	9,55	dpi
2	3	0,01	0,41	0,340	0,446	109,82	dpi/dbp
2	4	0,54	0,54	-0,102	0,103	0,01	ok
2	5	0,57	0,60	-0,075	0,128	0,35	ok
3	6	0,83	0,45	0,284	0,466	72,07	dpi
3	7	0,85	0,27	0,503	0,671	187,98	dpi/rediseño
3	8	0,47	0,56	-0,020	0,205	4,13	ok
4	9	0,16	0,34	0,090	0,264	19,48	dpi/dbp
4	10	0,27	0,24	-0,061	0,133	0,76	rediseño
4	11	0,11	0,12	-0,059	0,080	0,06	rediseño
5	12	0,27	0,44	0,071	0,270	16,36	dpi/dbp
5	13	0,69	0,46	0,132	0,336	30,11	dpi
5	14	0,29	0,40	0,007	0,208	6,54	dpi/dbp
6	15	0,65	0,41	0,139	0,355	30,89	dpi
6	16	0,39	0,30	-0,018	0,200	4,53	ok
6	17	0,43	0,38	-0,059	0,163	1,27	ok
7	18	0,59	0,57	-0,090	0,138	0,20	ok
7	19	0,43	0,41	-0,094	0,135	0,14	ok
7	20	0,77	0,59	0,083	0,286	17,24	dpi

pregunta, en total 20. Cada pregunta tiene un peso del 5% de la calificación máxima. Las dos siguientes columnas muestran los resultados por ítem de las EC para los Grupos B y A. En la quinta y sexta columna se muestran respectivamente los límites inferior y superior de un intervalo de confianza de la diferencia entre proporciones, al 99%. La siguiente columna muestra los valores obtenidos para Chi cuadrada y la última columna indica las decisiones tomadas en función de los resultados obtenidos.

dpi: Se tomó la decisión de descartar los ítems cuyo test de hipótesis resultara de rechazo, ello implica la pérdida de trazabilidad de las mediciones en las respuestas de estudiantes a los conceptos relacionados. Sin embargo, los MME encontrados siguen aportando conocimiento acerca de las dificultades de los estudiantes para comprender el modelo conceptual involucrado.

dpb: En aquellos ítems que además de rechazarse el test de hipótesis, se observaron valores muy bajos del estadístico de prueba correspondiente al Grupo B, se decidió retirar del banco de preguntas, la del ítem analizado. Ello en razón de que no es posible afirmar que la pregunta involucrada es representativa en su categoría.

rediseño: Esta decisión fue tomada cuando, a pesar de la aceptación del test de hipótesis, ambos estadísticos de prueba, tanto el perteneciente al Grupo B como el del A, arrojaban valores muy alejados del rendimiento académico medio.

ok: Cuando se acepta el test estadístico y el valor de la proporción es superior a 0,3, este último tomado como criterio de mínimo rendimiento admisible en función del histórico de la asignatura.

Sólo el 40% del total de preguntas superó el análisis de aptitud para posibilitar la continuidad de las acciones durante 2013.

Sin embargo, cuando se comenzó a implementar las acciones definidas en este análisis confirmatorio, los hallazgos observados obligaron a rectificar el rumbo una vez más.

Entre otros, y solo a modo de ejemplo, la pregunta correspondiente al ítem 3 del Grupo B se relaciona con un concepto distinto de la del Grupo A. Ello en razón de que en Moodle no se sorteó la del B al barajar las preguntas para el segundo grupo. Situaciones como éstas y similares, llevaron a la conclusión de la necesidad de rediseñar el banco de preguntas, acción que actualmente se está desarrollando. Se

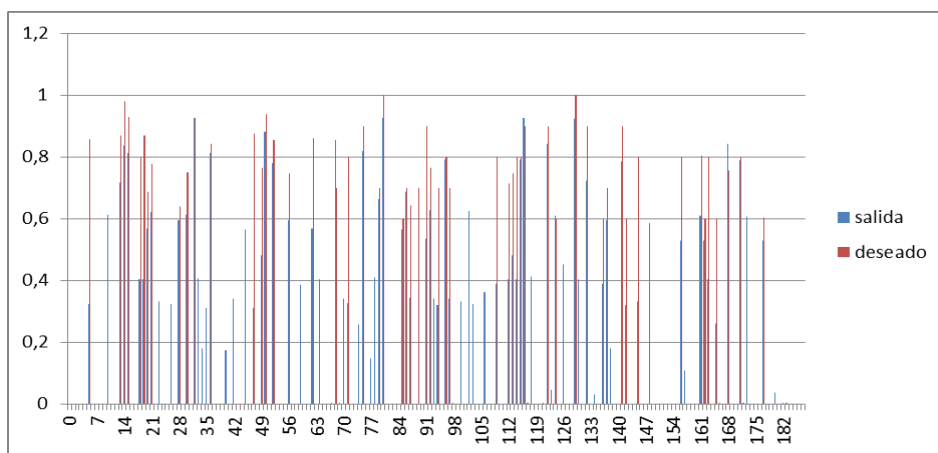
han propuesto un total de 70 ítems, que abarcan conceptos más acotados con lo cual la evaluación resultará más precisa y veraz.

Las preguntas relativas a los 8 ítems que superaron el análisis confirmatorio se preservan en la medida de lo posible en esta nueva organización para ser reusadas.

La información proveniente del análisis de los MME es utilizada en el nuevo diseño tanto para construir la base conocimiento de SIETTE como para la retroalimentación automática en Moodle.

En este próximo ciclo se realizarán experimentos con un Grupo testigo en SIETTE con el fin de avanzar sobre aspectos relativos al proceso de evaluación y sus resultados y además se implementará retroalimentación automática en la plataforma Moodle para las nuevas evaluaciones de opción múltiple, que se aplicará sobre el Grupo Control. Por otra parte, como las opciones de las evaluaciones de Moodle están preparándose en relación a los MME encontrados, sus resultados serán plausibles de alimentar Redes Neuronales Artificiales para detectar de patrones de aprendizaje de los estudiantes.

En la figura se muestran las predicciones obtenidas en la ejecución de la Red Neuronal para las Calificaciones de los alumnos 2012.



Se realizó análisis de capacidad de aprendizaje -comportamiento de la varianza en el tiempo- y exactitud de respuesta -diferencia entre los valores de salida y los esperados de los individuos-. No se determinaron intervalos de confianza para los distintos rangos de la variable de resultado.

Se tomaron los parámetros de ajuste de la red correspondientes a entrenamiento y validación con las muestras 2011-1 y 2010, respectivamente. La capa

oculta tiene ocho neuronas, el momentum nulo y delta de error generalizado bajo. La varianza promedio para las muestras utilizadas no superó el 0,8%.

La primera neurona de la capa de entrada representa el porcentaje obtenido en las Evaluaciones Conceptuales, las otras dos, que toman los valores: 2, 5 y 10, corresponden a los test predictores. La neurona de salida toma valores entre 0 y 100%, representando la calificación obtenida en la última instancia de Acreditación. Se normalizaron todos los datos.

Conclusiones

Si bien esta presentación muestra los avances hacia la detección de los modelos mentales de los estudiantes de programación; la conclusión más relevante tiene relación con la importancia de medir y comparar resultados tanto en la dimensión académica como en investigación.

De la interpretación del análisis estadístico se sugieren resultados superadores de los objetivos inicialmente planteados; ellos no sólo contribuyen a realizar ajustes y correcciones sobre el rumbo del proyecto de investigación, motivo de esta presentación.

Además nos proporcionaron una mirada crítica de nuestro desempeño como profesionales implementadores de estrategias de aprendizaje y evaluación. Si bien, esta asignatura se caracteriza por implementar año tras año diversas acciones de mejora tanto en los procesos de enseñanza como en los de evaluación, los análisis descritos en esta presentación posibilitan la revisión de estas acciones y la decisión de medidas correctivas en consonancia con ellas.

Con referencia a la ejecución de RNA, los resultados se comportan dentro de lo previsible, no habiendo encontrado hallazgos significativos que ameriten más comentarios. Se espera que la implementación de las RNA con aprendizaje no supervisado para detectar patrones de comportamiento de aprendizaje, hacia fines de 2013, resulte una contribución más alentadora.

Referencias

Bartó, C. A., and Díaz, L. C., 2012, "Intelligent Systems Applied to Computer Engineering Teaching". Revista IEEE Latin American Transactions.

Carretero, Mario y otros, 2005, *Introducción a la psicología cognitiva*. Aiqué Grupo Editor.

Díaz, L. C., Algorry, A. M, Eschoyez, M., Marangunic R. and Bartó, C. A., 2012. "Actions towards the application of intelligent systems in computer education". Revista IEEE Latin American Transactions

Johnson-Laird, Phillip N. Mental Models, 1993, *Deductive Reasoning, and the Brain*. Department of Psychoogy, University of Princeton. N. J. U.S.A.