Ingeniería de Software para clasificar patrones cognitivo conductuales

Marcelo Fabio Roldán[1], Germán Montejano[2], Ana Funes[2], Daniel Riesco[2]
[1]Departamento de Ciencias Exactas, Físicas y Naturales –
Universidad Nacional de La Rioja
La Rioja – Argentina
marcelofabio01@yahoo.com.ar
[2]Universidad Nacional de San Luis, Ejército de los Andes 950
5700 San Luis, Argentina
{gmonte, afunes, driesco}@unsl.edu.ar

Resumen

En la presente línea de investigación nos proponemos utilizar el Adaptive Business Intelligence (ABI) como herramienta de extracción de conocimiento en bases de datos, para clasificar probabilísticamente y con un grado de error aceptable, aquellas causalidades del fracaso de los aspirantes a sub oficiales de policía que ingresan en la escuela de la fuerza. A partir de allí, generar aplicaciones informáticas que permitan evaluar al futuro aspirante y determinar predictivamente su factibilidad de éxito en la escuela.

Esto es posible, en concordancia con la puesta en práctica de una metodología de reciente creación [RFM12, RDF13], la que facilita el desarrollo de aplicaciones basadas en Business Intelligence, favorece el uso de métodos de predicción y técnicas de optimización.

Este análisis de datos, para encontrar los patrones que determinan las causalidades de las bajas, lo realizaremos optimizando aquello que se denomina Extracción de Conocimiento de Bases de Datos (KDD), mediante una tecnología actualizada a través de la metodología que la implementa de manera simplificada y práctica. Estas causalidades principales, por las cuales los aspirantes no culminan su formación, se encontrarán de una manera probabilística.

Contexto

El presente trabajo de investigación se encuentra enmarcado en una colaboración entre investigadores del Proyecto de Incentivos "Ingeniería de Software para clasificar patrones cognitivo conductuales. Clasificación taxonómica predictiva y su impacto en la graduación de los aspirantes a la carrera de sub oficiales de policías", de la Universidad Nacional de La Rioja e investigadores del Proyecto de Incentivos "Ingeniería de Software: Conceptos, Métodos y Herramientas en un contexto de Ingeniería de Software en Evolución", de la Universidad Nacional de San Luis.

Este trabajo de investigación surge de la aplicación de la metodología presentada en su tesis de Maestría en "Ingeniería de Software" en la Universidad Nacional de San Luis por uno de los autores de este trabajo [RFM12, RDF13].

Palabras clave: Data Mining, Adaptive Business Intelligence, Cognitivo Conductual, Psicometría.

Introducción

El Business Intelligence se nutre de grandes volúmenes de datos para permitir que los usuarios y aplicaciones accedan a ellos y los usen para realizar toma de decisiones

[WAT07] [AYA06] [REI00]. Es así que metodologías como CRISP-DM [CHA00] [GUT07] [AZE08], SEMMA [SAS11] [TAA11] y, en general, el KDD Process, nos introducen al proceso utilizado para alcanzar tal objetivo por medio del uso de técnicas de Minería de Datos. En particular, la metodología ABI a seguir [RFM12, RDF13], desde un punto de vista integrador y a partir de metodologías de data mining comunes, propone una serie de pasos a seguir, necesarios para alcanzar los resultados esperados. Su aplicación para el análisis prospectivo, avanza mucho más allá del razonamiento basado en eventos pasados que proveen las herramientas típicas de los sistemas de soporte a las decisiones. El conocimiento obtenido se potencia al aplicar esta metodología, pudiendo responder preguntas de dificultosa resolución o que demandan excesivo tiempo.

Con esta tecnología se pretende encontrar las causalidades para su aplicación predictiva probabilística en nuevos casos. Como lo indica Deshpande [DES10] el data mining predice tendencias futuras y eventos, ayuda a las organizaciones a la toma de decisiones proactivas basadas en conocimiento.

Teniendo en cuenta lo dicho anteriormente, creemos que los algoritmos de minería de datos se pueden usar para reducir los costos en la academia de policía de manera significativa.

Este uso de minería de datos en la determinación de perfiles psicológicos adecuados, es una aplicación con gran potencial, ya que la información psicológica es compleja y difícil de analizar.

Resultados y Objetivos

Este trabajo de investigación tiene como objetivo principal aplicar una metodología, basada en las estructuras conceptuales de Adaptative Business Intelligence a un caso particular que sirva de referencia para su aplicación posterior a otros ámbitos de las fuerzas de seguridad.

Esperamos que mediante el uso y las aplicaciones de Adaptive Business Intelligence en contextos predictivos, nos sea posible determinar probabilísticamente y con un grado de error aceptable, aquellas causalidades del fracaso de los aspirantes a policías que ingresan en la escuela de la fuerza tomando como premisa los patrones cognitivo conductuales y su clasificación taxonómica.

El trabajo lo llevaremos a cabo a partir de casos de estudio, y de bibliografía de referencia que permitan facilitar el uso de métodos de predicción.

Durante el proyecto se prevé utilizar las herramientas de Inteligencia Artificial avanzada siguiendo los lineamientos establecidos en la metodología mencionada, para encontrar respuestas a una problemática que se presenta en la instancia preliminar de formación de la policía en su escuela de sub-oficiales. Este problema se manifiesta como fracaso en la continuidad de los estudios que los policías deben aprobar para alcanzar la condición de tales. Los datos relacionados con el problema nos permitirán constatar que el fracaso de los estudios en las escuelas de policías, es un fenómeno con múltiples orígenes, tanto sociales, culturales como del perfil psicológico de los aspirantes o de sus aspectos académicos. El resultado de la clasificación resulta en dos clases: “Apto” y “No Apto”, concordando con la evaluación que el profesional del área de Psicología Organizacional determina de manera analítica en cada ficha de aspirante.

Ante este escenario, la aplicación de la nueva metodología ABI, facilitará la producción de conocimiento a partir del aprendizaje y del descubrimiento.

Desarrollaremos durante el proyecto, las actividades necesarias para normalizar los datos de tests psicométricos, pertenecientes a historiales existentes de los policías aspirantes a la academia de sub oficiales de

policía de La Rioja. Posteriormente, sobre estos casos aplicaremos la Metodología Adaptive Business Intelligence en contextos Predictivos [RFM12, RDF13]. A través de la aplicación de las tecnológicas informáticas analizaremos el problema particular del fracaso de los ingresantes a la escuela de sub oficiales buscando encontrar patrones asociados a la misma.

Para el desarrollo de la aplicación predictiva, en este caso, trabajaremos con datos de aspirantes de la escuela de policía de La Rioja en el período que va del año 2005 al año 2010. Estos datos corresponden a personas reales, de las cuales se preserva su identidad, conservados en un archivo profesional cuyo acceso está restringido a estos estudios.

En una etapa posterior, correlacionaremos los resultados arrojados por los instrumentos de evaluación psicométrica con los perfiles de aptitud de los sub oficiales de policía. Estos resultados nos servirán asimismo para analizar la calidad de predicción de los modelos generados [PAL05].

Buscamos de esta forma facilitar el desarrollo informático y, mediante la aplicación de esta metodología en un trabajo interdisciplinario, demostrar su eficacia en un marco orientado a herramientas de extracción de conocimiento, usando problemas reales como eje ilustrativo de su aplicación.

En relación a los aspectos psicométricos sobre los cuales se han realizados las mediciones a los aspirantes – sujetos de las pruebas que son fuente de información para el proyecto- existe una amplia bibliografía al respecto, cuyo basamento científico está ampliamente demostrado [CUE93].

Se conoce que el índice de bajas figura como indicador de calidad y tiene una gran incidencia en el costo operativo del funcionamiento de las fuerzas policiales, ya que de la masa inicial que ingresa a las filas como aspirantes, un reducido porcentaje logra alcanzar la finalización exitosamente; de allí la importancia de contar con indicadores para una detección precoz.

Líneas de Investigación y Desarrollo

La línea de investigación aquí presentada, la cual está siendo llevada adelante entre investigadores del proyecto de incentivos de la Universidad Nacional de La Rioja, “Ingeniería de Software para clasificar patrones cognitivo conductuales. Clasificación taxonómica predictiva y su impacto en la graduación de los aspirantes a la carrera de sub oficiales de policías.” e investigadores del Proyecto de Incentivos “Ingeniería de Software: Conceptos, Métodos y Herramientas en un contexto de Ingeniería de Software en Evolución”, de la Universidad Nacional de San Luis, se trata de la continuación de una colaboración existente entre ambas instituciones. Dicha línea de investigación surge de la aplicación de la metodología ABI [RFM12, RDF13], la cual fue desarrollada como tesis de Maestría en Ingeniería de Software de la Universidad Nacional de San Luis por uno de los autores de este trabajo.

Formación de Recursos Humanos

El presente trabajo surge como una de las tantas aplicaciones posibles de la metodología desarrollada en un trabajo de tesis de posgrado (Maestría en Ingeniería de Software de la Universidad, Nacional de San Luis). La misma ha sido recientemente defendida, y ha facilitado el desarrollo de proyectos de investigación entre investigadores de la Universidad Nacional de la Rioja y de la Universidad Nacional de San Luis, bajo un acuerdo entre ambas universidades.

Referencias

[ARA11] J. G. Arancibia. Metodología para el Desarrollo de Proyectos en Minería de Datos CRISP-DM.
URL: http://yoshibauco.wordpress.com/ - Último acceso: 2011

[ASE04] V. V. Asencios. Data Mining y el Descubrimiento del Conocimiento. Revista de la Facultad de Ingeniería Industrial. Vol (7). 2004

[AYA06] Alejandro Peña Ayala. Inteligencia de Negocios: Una Propuesta para su desarrollo en las organizaciones. Instituto Politécnico Nacional. Dirección de Publicaciones. 2006

[AZE08] A. Azevedo. M. F. Santos. KDD, SEMMA y CRISP-DM: A Parallel Overview. IADIS European Conference Data Mining 2008 (part of MCCSIS 2008)

[CHA00] Pete Chapman, Julian Clinton, Randy Kerber, Thomas Khabaza, Thomas Reinartz, Colin Shearer and Rüdiger Wirth. CRISP-DM 1.0 Step-by-step data mining guide, NCR Systems Engineering Copenhagen (USA and Denmark), DaimlerChrysler AG (Germany), SPSS Inc. (USA) and OHRA Verzekeringen en Bank Groep B.V (The Netherlands). Copyright © 1999, 2000

[CIO07] K.J.Cios,W.Pedrycz, R.W.Swiniarski, L.A.Kurgan. Data Mining. A Knowledge Discovery Approach. Springer. 2007

[CUE93] Cueto, E. G., Introducción a la psicometría. Editores: Madrid: España, 1993

[DES10] Data Mining System and Applications: A review. S.P.Deshpande. V.M. Thakare. International Journal of Distributed and Parallel systems. Vol 1. No 1. 2010

[GUT07] P-Chapman. J. Clinton. R. Kerber. T. Khabaza. T. Reinartz. C. Shearer. R. Wirth.
Guía paso a paso de Minería de Datos.

[HAL09] Mark Hall, E. Frank, G. Holmes, B. Pfahringer. The WEKA Data Mining Software: An Update. Pentaho Corporation. 2009

[IEE07] Watson, H.J.; Wixom, B.H. The Current state of the Business Intelligence.
Univ. of Georgia, Athens. 2007. Último acceso 2011 http://ieeexplore.ieee.org/

[MIC07] Z. Michalewicz, M. Schmidt, M. Michaelewicz, C. Chiriac. Adaptative Business Intelligence. Springer 2007

[MOS03] Larissa T. Moss, Shaku Atre. Business Intelligence Roadmap: The Complete Project Lifecycle for Decision-Support Applications. Addison Wesley. 2003

[PAL05] Pal, N.R., Jain, L.C. Advanced Techniques in Knowledge Discovery and Data Mining, Springer Verlag. (Eds.) 2005.

[REI00] Joerg Reinschmidt, Allison Francoise. Business Intelligence Certification Guide IBM. International Technical Support Organization. 2000

[RMF12] Roldán, M. F., Montejano, G., Funes, A, Una Metodología para el Desarrollo de Aplicaciones Autoadaptativas basada en Business Intelligence. Aplicación en Medicina. Tesis de Maestría. 2012

[RDF13] Roldán, M. F., Debnath, N., Funes, A, Montejano, G., Riesco, D. A Methodology Based on Business Intelligence for the Development of Predictive Applications in Self-Adapting Environments, a aparecer en proceedings de ICCSEE 2013, Hangzhou,China.

[WAT07] Hugh J. Watson. Barbara H. Wixom. The Current State of Business Intelligence. University of Georgia. University of Virginia. 2007. IEEE Computer Society http://www.computer.org

[ZBI06] Adaptive Business Intelligence , Zbigniew Michalewicz, Springer 2006