

Inteligencia Artificial aplicada al desarrollo de Evaluaciones de Matemática

Gabriel Sosa, Lucas Dima, Rafael Urdaneta, Gabriela Esperón, Daniela López De Luise.

AIGroup Facultad de Ingeniería. Universidad de Palermo. Mario Bravo 1050 8º piso.
Buenos Aires. Argentina.
aigroup@palermo.edu

Resumen. El presente trabajo expone una propuesta de arquitectura, que emplea herramientas de minería de datos y un sistema experto, para detectar factores socioeconómicos y contextuales que podrían interferir en el proceso de aprendizaje de alumnos universitarios en el área de Matemática.

Para ello se releva un conjunto de variables relacionadas con los factores mencionados de la población de interés y se implementa una evaluación prediseñada por docentes especializados. Los datos así obtenidos y las respuestas de los alumnos a la evaluación propuesta son analizados con algoritmos exploratorios para obtener características por grupos.

Se emplea el método jerárquico de Ward para determinar el número óptimo de clusters y luego el procedimiento k-medias para formar conglomerados. Dichos algoritmos generarán las métricas para la catalogación de los datos. Este agrupamiento constituirá la base para la formulación de reglas de un sistema experto que generará recomendaciones a sus usuarios.

Palabras clave: evaluación, sistemas expertos, detección de problemas en el aprendizaje, minería de datos, modelos, arquitectura, algoritmos de agrupamiento.

1 Introducción

Hoy en día los especialistas en educación ponen de relieve las múltiples acepciones, objetivos e implicancias del término evaluación. El aspecto de la evaluación que interesa en el presente trabajo es aquel relacionado con la evaluación de los aprendizajes de los alumnos.

Evaluar supone conocer qué y para qué se va a evaluar. Un requisito esencial para ello es recoger información, analizar y valorar las características y condiciones de un alumno, un grupo de estudiantes, o un entorno educativo en función de parámetros de referencia con el fin de tomar decisiones con vistas al futuro, tendientes a mejorar los procesos de aprendizaje.

Si se considera la dimensión formativa que debería tener el proceso de evaluación, resulta de fundamental importancia aludir al concepto de "retroalimentación", la cual

supone la identificación, en mayor o menor medida, de las fortalezas y debilidades en el conocimiento y habilidades de los estudiantes, así como la identificación de las causas que dificultan dicho proceso. [1]

Es aquí donde distintas técnicas de inteligencia artificial tales como algoritmos de agrupamiento y sistemas expertos podrían constituirse en poderosas herramientas tanto para identificar posibles causas vinculadas a un deficiente progreso en el desempeño académico y el proceso de aprendizaje, como para proveer una retroalimentación adecuada que a partir de recomendaciones efectuadas por el sistema permita a los alumnos revisar y corregir sus propios procesos. Por ejemplo, los sistemas expertos ofrecen la posibilidad de diagnosticar, depurar y corregir las producciones de los estudiantes en un área particular del conocimiento ya que permiten determinar el nivel cognoscitivo de los alumnos y ayudar a desarrollar y mejorar sus habilidades.[2]

Adicionalmente, la implementación de un sistema como el proyecto Einstein, que aquí se propone, también brindaría información a los docentes para el diseño de situaciones didácticas que permitan mejorar la calidad del aprendizaje [3]. Este proyecto se desarrolla en el marco del AIGroup de la Universidad de Palermo, en conjunto con el Instituto Tecnológico de Monterrey, México.

El resto del trabajo se organiza como sigue: la sección 2 describe la arquitectura del sistema; en la sección 3 se describe cada uno de los módulos componentes del mismo; en la sección 4 se presentan las conclusiones y trabajo a futuro.

2 Arquitectura

El sistema está siendo desarrollado en el lenguaje orientado a objetos Java versión 1.5 (TM) [4]. La elección del mismo se debe al alto grado de portabilidad, es decir que sus aplicaciones son independientes del sistema operativo en el que se ejecuten. Otras características importantes son su madurez, la gran cantidad de herramientas ya desarrolladas y de aportes existentes por parte de la comunidad.

El sistema toma como entrada los datos de un grupo de estudiantes junto con la definición de los mismos. Esta definición consta de dos grupos: definición de la estructura de las respuestas (tabla 1) y definición de datos socioeconómicos (tabla 2).

Tabla 1. Definición de la estructura de las respuestas.

Variable	Tipo	Utilización	Ejemplo
Identificador único	Alfanumérico	Identificador interno del sistema	PRAX412
Nombre de la clase de la respuesta	Alfanumérico	Corresponde al nombre completo de clase Java que identifica al tipo de respuesta	<i>edu.palermo.einstein.answers.MultipleOptionNotOrderedAnswer</i>
Estrategia de análisis de la respuesta [5]	Alfanumérico	Indica la estrategia utilizada para realizar el análisis de los datos. Corresponde al nombre completo de clase Java	<i>edu.palermo.einstein.parser.strategy.CsvParserStrategy</i>

Tabla 2. Definición datos socioeconómicos.

Variable	Tipo de dato	Ejemplo
Identificador único	Alfanumérico	UPAL2001004128
Fecha de nacimiento	Fecha	1983-05-18
Recurrente de la asignatura	Booleano	Falso
Casado	Booleano	Verdadero
Trabaja	Booleano	Verdadero
Cantidad de hijos	Númérico	0
Quien paga los estudios	Lista de opciones	Uno mismo
Relación con la profesión	Booleano	Verdadero
Cantidad de horas diarias de trabajo	Númérico	8
Cantidad de horas semanales de estudio	Númérico	40
Viene de otra universidad	Booleano	Falso

Los datos y la definición de los mismos son ingresados al módulo de interpretación el cual genera los objetos que representan a un individuo y que contiene sus datos socioeconómicos tanto como las respuestas a la evaluación realizada. Estas representaciones se van almacenando en una base de datos orientada a objetos para el posterior análisis por el módulo inteligente. Este último es el encargado de implementar herramientas de minería de datos, mediante las cuales:

- Se accede a los objetos anteriormente mencionados.
- Se aplican los algoritmos de agrupamiento y se obtienen los conglomerados.
- Se formulan las reglas del sistema experto en función de las características predominantes en cada uno de los conglomerados.
- Las reglas permitirán que el sistema brinde recomendaciones a los usuarios acerca de las habilidades matemáticas que se deberían revisar con el fin de afianzarlas, modificarlas o mejorarlas.
- El módulo inteligente realizará la devolución de los resultados al usuario a través de una interfaz gráfica a la vez que almacenará nuevamente los datos en el Almacenamiento Persistente.

La figura 1 muestra el diagrama de proceso.

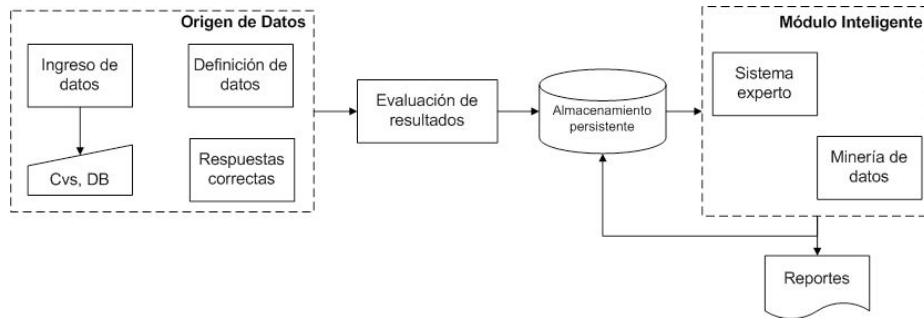


Fig. 1. Diagrama de proceso.

El funcionamiento del sistema en orden cronológico se puede ver en el siguiente diagrama de secuencia (figura 2)

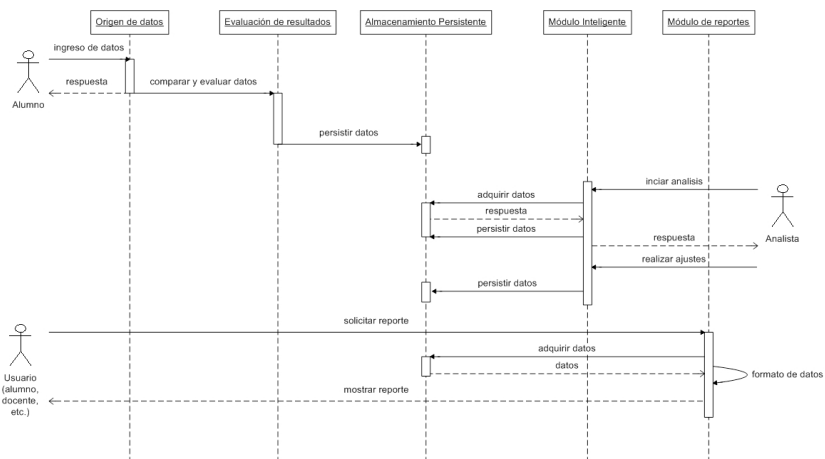


Fig. 2. Diagrama de secuencia.

3 Descripción de los Componentes

3.1 Módulo de Origen de Datos

El módulo Origen de datos está formado por tres componentes: Ingreso de Datos, Definición de Datos y Respuestas Correctas.

Ingreso De Datos

El sistema toma como entrada los datos de un grupo de estudiantes de una asignatura dada y las respuestas de un conjunto de ejercicios. Este proceso consta de dos partes. La primera es la carga de los datos socioeconómicos estadísticos de cada una de las personas pertenecientes al grupo a evaluar y la segunda es la resolución del conjunto de ejercicios.

Para los datos socioeconómicos se escogieron una serie de variables relevantes que luego serán evaluadas por el resto de los componentes del sistema. La figura 3 muestra el módulo de ingreso de datos.

The screenshot displays the 'Einstein' data entry module. At the top, there's a header with the name 'Einstein' and a 'Descripcion' label. Below this is a progress bar showing 0% completion. A section titled 'Datos personales' states 'Todos los datos son anónimos'. The form is divided into two columns. The left column contains fields for 'Legajo' (a text input with a note 'Solo se aceptan números en este campo'), 'Sexo' (radio buttons for 'Femenino' and 'Masculino'), 'Fecha de Nacimiento' (a date picker with the format 'dd.mm.yyyy'), and '¿Es recursante de esta materia?' (radio buttons for 'Si' and 'No'). The right column contains fields for '¿Viene de otra Universidad?' (radio buttons for 'Si' and 'No'), '¿Es casado?' (radio buttons for 'Si' and 'No'), '¿Trabaja?' (radio buttons for 'Si' and 'No'), '¿Cantidad de hijos?' (a dropdown menu with the text 'Seleccione una de las siguientes opciones'), and a final dropdown menu labeled 'Seleccione, por favor...'.

Fig. 3. Módulo de Ingreso de datos

La segunda parte de los datos de entrada corresponde a las respuestas de un conjunto de ejercicios por parte cada uno de los individuos. En cuanto a la parte técnica y en virtud de mantener todo el proyecto utilizando software código abierto (open source) con licenciamiento GNU GPL se eligió para la carga de datos la aplicación web Limesurvey [6], que es completamente configurable y posee una interfaz clara y concisa, indispensable para que el ingreso de datos no entorpezca el proceso, y sea lo más sencillo y rápido posible para el individuo a examinar. Se

calcula que con un conjunto de ejercicios estándar cada individuo puede tardar entre 10 y 15 minutos para cargar sus datos y sus respuestas.

El hecho de que la carga de datos se realice a través de una aplicación web permite que se puedan evaluar muestras independientemente de su ubicación geográfica, logrando un buen complemento para entornos de educación a distancia.

Junto con los datos de entrada, se incluirá una clave única por cada alumno que permitirá la realización de un seguimiento longitudinal del desempeño de este individuo.

Uno de los puntos clave de la arquitectura del proyecto es que permite interactuar con datos en cualquier formato. Para lo cual se decidió que los mismos (los datos) se encuentren separados de sus meta-datos [7].

Definición de Datos

Este componente es el encargado de interpretar la definición de los datos propiamente dichos y lograr que el sistema vea de manera transparente los datos que requiere procesar. La definición de los datos se encuentra en un archivo modelado en formato YAML [8], el cual contiene lo necesario para interpretar los datos de manera adecuada. Se optó por este formato ya que el mismo se encuentra ampliamente adoptado y es de fácil interpretación para el operador del sistema. Este enfoque permitiría que la definición de los datos pueda ser modificada sin realizar cambios en el núcleo del sistema, ya que está almacenado de forma externa en un archivo de texto plano. También permitiría un ágil manejo por parte del administrador y le otorgaría flexibilidad y versatilidad a todo el sistema.

Respuestas Correctas

Contiene las respuestas correctas que se utilizarán para contrastar los resultados de los ejercicios que realizan los alumnos.

3.2 Módulo de Evaluación de Resultados

Este módulo recibe los archivos procesados por el Módulo de Origen de datos y se encarga de realizar la comparación entre las respuestas generadas por los estudiantes y el conjunto de respuestas correctas.

En esta primera etapa se han considerado tres tipos de respuestas para los ítems:

- *Respuesta simple cerrada*: donde el estudiante debe seleccionar una opción dentro de las listadas.
- *Respuesta múltiple no ordenada*: permite seleccionar varias opciones entre las listadas y donde el orden dado de las respuestas no es considerado.
- *Respuesta múltiple ordenada*: permite seleccionar varias respuestas y debe tener en consideración el orden dado de las respuestas.

En este punto los registros de los estudiantes se encuentran en un formato abstracto, conteniendo la información socioeconómica de cada uno de ellos, como así también las respuestas de cada uno de los ítems de los ejercicios.

Los registros serán almacenados para luego ser procesados por el Módulo Inteligente.

3.3 Módulo Inteligente

Se encuentra dividido en dos componentes.

El componente de Minería de Datos obtendrá los datos de los resultados de las evaluaciones procesadas y persistidas por el módulo de Evaluación de Resultados (Parser), seguidamente ejecutará un método jerárquico de Clustering para definir el número óptimo de conglomerados.

Los métodos jerárquicos se emplean mediante un proceso de unión o división sucesiva de los elementos. El resultado final es una jerarquía de unión, que se representa normalmente en forma de árbol denominado dendograma (figura 4) en la que cada elemento se une o separa en una determinada fase del proceso.

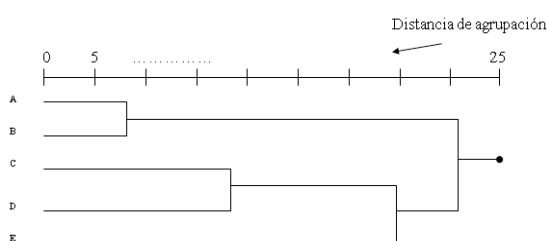


Fig. 4. Dendograma

Ward (1963), propuso un método de cluster jerárquico muy general, conocido como “Método Ward” o el “Método de la varianza mínima”. La propuesta de este método se basa en que la pérdida de información que se produce al integrar los distintos individuos en clusters puede medirse a través de la suma total de los cuadrados de las desviaciones entre cada punto (individuo) y la media del cluster en el que se integra. Para que el proceso de agrupamiento resulte óptimo, en el sentido de que los grupos formados no distorsionen los datos originales, Ward propuso la siguiente estrategia:

En cada paso del análisis, considerar la posibilidad de la unión de cada par de grupos y optar por la fusión de aquellos dos grupos que menos incrementen la suma de los cuadrados de las desviaciones al unirse.

Si bien es cierto que no existen criterios objetivos y ampliamente válidos para la determinación del número óptimo de clusters, nos apoyaremos en la observación de la caída brusca en la similitud de los elementos a partir de los datos producidos empleando el método de ward (figura 5), dado que este método es uno de los más empleados para este objetivo; posee casi todas las ventajas del método de la media y suele ser más discriminativo en la determinación de los niveles de agrupación. Una investigación llevada a cabo por Kuiper y Fisher probó que este método era capaz de acertar mejor con la clasificación óptima que otros métodos (mínimo, máximo, media y centroide).[9]

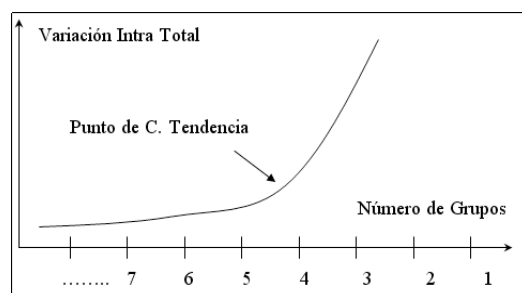


Fig. 5. Variación Intra Total

Luego de determinar el número óptimo de cluster, se continúa mediante el procedimiento de k-medias para formar los grupos [10].

En este componente se integrarán las respuestas de los alumnos a los ejercicios y sus correspondientes datos socioeconómicos. El estudio exhaustivo de las características de los conglomerados permitirá la elaboración de perfiles que posibilitarán la creación de las reglas que empleará el Sistema Experto [11] [12].

Cada grupo generado, representa a un conjunto de alumnos con características similares y a su vez cada grupo se distingue de otro por su heterogeneidad entre ellos. De esta manera obtenemos grupos con características particulares en base a las variables definidas en el primer módulo del sistema. A efectos de los estudios realizados y en base a los resultados obtenidos, se podrán ajustar los tipos y cantidades de variables en la medida de lo necesario para producir y comparar resultados alternos.

Este módulo está construido utilizando las librerías del motor de Weka (Waikato Environment for Knowledge Analysis - Entorno para Análisis del Conocimiento de la Universidad de Waikato) [13], el cual es un conocido software para aprendizaje automático y minería de datos escrito en Java. WEKA es un software libre distribuido bajo licencia GNU GPL y dado que Einstein se está desarrollando totalmente sobre el lenguaje Java, se logra una integración natural y transparente en el producto final, pudiendo ejecutarse en cualquier plataforma que tenga el JRE de Java (Java Runtime Environment) versión 1.5 o superior.

El segundo componente de este módulo es el Sistema Experto, el cual es el encargado de realizar la evaluación final de los grupos significativos revelados por el componente de Minería de Datos.

Será construido utilizando Drools, un sistema de administración de reglas de negocio (BRMS) con un motor de reglas basado en una adaptación orientada a objetos del algoritmo Rete [14].

Drools permite expresar de una forma más natural las reglas de diagnóstico, interactuando con los objetos propios del sistema. Provee separación de lógica (reglas) y datos (hechos). También aporta soporte para la programación declarativa, y es lo suficientemente flexible para expresar la semántica del problema con un lenguaje específico de dominio (DSL). Drools utiliza el lenguaje de reglas de drools (DRL) para especificar las condiciones, acciones y funciones de las mismas, las cuales se pueden expresar con distintos lenguajes, como Java y MVEL. Luego las reglas serán guardadas en archivos de texto con la extensión drl.

De esta manera las reglas de diagnóstico, que normalmente son muy cambiantes, podrán ser ajustadas de una manera externa al sistema sin necesidad de rediseñar, modificar y redistribuir el sistema [15].

3.4 Módulo de Reportes

En el se desplegarán los resultados generados a partir de la aplicación de las reglas. Se visualizarán las recomendaciones para un grupo de individuos, para un individuo en particular y en forma optativa las reglas aplicadas.

Se visualizarán los parámetros del agrupamiento generado por el algoritmo de Minería de Datos, como características socioeconómicas de cada grupo en relación con los resultados de la evaluación.

Individualmente, se podrán ver las respuestas generadas y su comparación con las respuestas correctas de un estudiante en particular, junto con sus datos socioeconómicos.

En virtud de la portabilidad de los reportes generados, el módulo permitirá utilizar diversos formatos estándar.

4 Conclusión

Se ha finalizado el desarrollo de los tres componentes del Módulo Origen de Datos, se ha implementado una primera evaluación consistente en una serie de problemas referidos a temas de Álgebra Lineal y se han realizado los ajustes para iniciar la captura de datos.

Se espera que la fase de captura de datos y el posterior análisis por el componente Minería de Datos del Módulo Inteligente permita, no sólo detectar patrones y construir los perfiles de alumnos con características similares, necesarios para elaborar las reglas del Sistema Experto, sino que brinde la posibilidad de poner a prueba el Módulo Origen de Datos y realizar los ajustes y correcciones necesarios en los componentes de dicho módulo. También que permita conocer si el tipo de respuestas incorporadas en el módulo Evaluación de Resultados es la más adecuada para los objetivos del proyecto o si resulta necesario incorporar otras alternativas. Asimismo que brinde información acerca de la pertinencia de los ejercicios planteados en esta primera evaluación en función de los objetivos del proyecto.

A partir de la recopilación y análisis de datos se espera poder realizar una primera aproximación que permita tipificar los errores en las producciones de los alumnos lo cual serviría en primer lugar para desarrollar las futuras evaluaciones y, en segundo lugar, permitiría ir perfeccionando progresivamente el sistema en función de los resultados que se vayan obteniendo.

Queda pendiente el desarrollo de evaluaciones dinámicas que, a partir de la modificación de distintos parámetros, permitan la creación de diversos conjuntos de problemas y ejercicios para un mismo contenido.

En las próximas etapas se irán diseñando e incorporando nuevos conjuntos de actividades para la evaluación de los distintos contenidos de la asignatura en la que se está implementando el proyecto.

En la fase siguiente está previsto aplicar el sistema tanto a la evaluación de otros contenidos del área de Matemática como a otras asignaturas, ya que una de las características principales del proyecto es la posibilidad de transferencia a otros campos gracias a su arquitectura modular que permite extender las funcionalidades del sistema sin modificaciones sustanciales en el mismo. Esta última característica otorgaría una gran versatilidad al sistema.

Referencias

1. Camilloni A. et al, ¿Qué significa evaluar? En La evaluación de los aprendizajes en el debate didáctico contemporáneo, pp. 35-66. Paidós (1998)
2. Vélchez Quesada, E.:Sistemas Expertos para la enseñanza y el aprendizaje de la Matemática en la Educación Superior. CIEMAC (2007)
3. Theory of didactical situations in Mathematics.G. Brousseau.pp 30-31
4. Sun Trademark and Logo Usage Requirements, <http://www.sun.com/policies/trademarks>
5. Erich Gamma, et al. "Design Patterns: Elements of Reusable Object-Oriented Software". pp 315. Addison-Wesley. (1995)
6. Lime Survey, <http://www.limesurvey.org/>
7. Senso, J.A., El concepto de metadato. Algo más que descripción de recursos electrónicos, Ci. Inf., 32, pp. 95-106 (2003)
8. YAML: a human friendly data serialization standard for all programming languages, <http://www.yaml.org/>
9. Johnson, R., Wichern, D., Applied Multivariate Statistical Analysis,pp. 690-691, Prentice Hall (2002)
10. Witten I., Frank E, Data Mining: Practical Machine Learning Tools and Techniques, pp. 265, Morgan Kaufmann (2005)
11. King G., Quantitative Discovery of Qualitative Information: A General Purpose Document Clustering Methodology, Talk at Washington University, St. Louis (2010)
12. UCLA Academic Technology Service, <http://www.ats.ucla.edu/stat/spss/>
13. WEKA, <http://www.cs.waikato.ac.nz/ml/weka/>
14. Forgy C., Rete: A Fast Algorithm for the Many Pattern/Many Object Pattern Match Problem, Artificial Intelligence, 19 (1982)
15. Drools, <http://www.jboss.org/drools/documentation.html>