



GESTÃO UNIVERSITÁRIA NA AMÉRICA DO SUL I COLÓQUIO INTERNACIONAL

25, 26 e 27 de outubro de 2000

Florianópolis, Santa Catarina - Brasil

COMUNICAÇÕES CIENTÍFICAS

ÁREA 4: ESTRUTURA DO ENSINO SUPERIOR

LAS REDES NEURONALES ARTIFICIALES COMO ELEMENTOS DEL SISTEMA DE SOPORTE A LAS DECISIONES EN LA ADMINISTRACIÓN UNIVERSITARIA

Lucía I. Passoni*

Resumen: Se propone a las redes neuronales como herramientas del sistema de soporte a las decisiones dentro del ámbito de la gestión académica. La experiencia se realiza en el ámbito académico de una institución universitaria, en la Facultad de Ingeniería de la Universidad Nacional de Mar del Plata. Se elige a los mapas auto-organizativos basados en el algoritmo de Kohonen. Los mismos constituyen un tipo de red neuronal artificial de aprendizaje no supervisado. Con el fin de validar la propuesta se lleva a cabo un experimento sobre una base de datos académicos. Específicamente se aplica sobre los datos de alumnos de la carrera de Ingeniería Electrónica de la Facultad de Ingeniería de la Universidad Nacional de Mar del Plata. La aplicación de este proceso genera el agrupamiento topológico en un mapa bidimensional de los datos multivariados de la historia académica de los alumnos, permitiendo la identificación de características de interés para el gerenciamiento en el ámbito departamental. Como resultado de la experiencia se ha logrado extraer conocimiento interesante. A partir de una base de datos del tipo puramente cuantitativa se han observado situaciones de interés para la toma de decisiones del área académica, como es: *el grado de avance relativo junto con la calificación de una cohorte de alumnos desagregada por áreas del conocimiento*. Se llega a la detección de áreas con identificación que influyen fuertemente en el avance de las cohortes, el conocimiento de esta situación facilita el proceso decisorio cuando se discute la reforma curricular metodológica y de contenidos.

Palabras clave: Sistemas de Soporte a las Decisiones; Redes Neuronales Artificiales; Minería de Datos; Gestión Universitaria.

*Lucía Isabel Passoni

Ingeniera electricista con orientación electrónica.

Profesor adjunto de Ingeniería Electrónica

Miembro del equipo de investigadores del Laboratorio de Bioingeniería

Facultad de Ingeniería

Universidad Nacional de Mar del Plata

Dirección postal: Laboratorio de Bioingeniería

Facultad de Ingeniería. Universidad Nacional de Mar del Plata.

Juan B. Justo 4302

Mar del Plata (7600)

Argentina

e-mail: lpassoni@fi.mdp.edu.ar

1. Introducción

Las nuevas tendencias en administración de instituciones de educación superior han propiciado la incorporación de técnicas de gestión “importadas” del mundo empresarial con el fin de elevar la eficiencia y la efectividad de su desempeño. Es así como nuevas tecnologías organizacionales tales como la calidad total, la planificación estratégica y la gerencia por objetivos han sido incorporadas en forma reciente en las instituciones universitarias norteamericanas¹. Si bien existen ciertos factores de especificidad de la gestión pública que diferencia a la misma del mundo empresarial, sin lugar a dudas las técnicas gerenciales aportan herramientas adecuadas al reto de la modernización de las administraciones públicas². El modelo de decisión racional que subyace en las técnicas de gestión empresarial sitúa su ámbito de aplicación en la actuación de las organizaciones públicas individualmente consideradas y en su funcionamiento interno.

Esta problemática se evidencia en las universidades cuando se intenta abordar cuestiones relacionadas esencialmente con el financiamiento y la planificación. El diseño del presupuesto institucional, la presentación de proyectos con fines específicos, la creación de nuevas carreras, son decisiones que deben ser sustentadas en el conocimiento del estado del sistema.

A partir de las últimas tendencias en la ciencia de la administración la gestión universitaria puede ser abordada desde la Gerencia del Conocimiento (*Knowledge Managment*). Este es uno de los temas cruciales del proceso de administración dentro de las organizaciones y una fuente principal de ventaja competitiva³.

Teóricos del gerenciamiento tales como Peter Drucker⁴, y Peter Senge⁵ han contribuido a la evolución del Gerenciamiento del Conocimiento. Drucker ha enfatizado la creciente importancia de la información y del conocimiento explícito como recursos

¹ FANELLI, A. **Gestión de las Universidades Públicas**: la experiencia internacional. Buenos Aires, Argentina: Ministerio de Cultura y Educación, 1998.

² ANDREU, R.; RICARTE, J. **Valor**: Estrategia y Sistemas de Información. Madrid: McGraw-Hill, 1994.

³ The Biztech Network <http://www.brint.com> Knowledge Management & Organizational Learning, 1999.

⁴ DRUCKER, P. **La Sociedad Poscapitalista**. Buenos Aires: Edit. Sudamericana, 1993.

⁵ SENGE, P. **The Fifth Discipline: the arte and practice of the learning organization**. NY, 1990.



GESTÃO UNIVERSITÁRIA NA AMÉRICA DO SUL I COLÓQUIO INTERNACIONAL

25, 26 e 27 de outubro de 2000

Florianópolis, Santa Catarina - Brasil

organizacionales, mientras que Senge ha abordado una dimensión cultural de la gerencia, considerando a las empresas e instituciones como **organizaciones que aprenden** (*learning organizations*).

La GerenciaC del onocimiento es un dominio que atraviesa disciplinas entre las que se encuentran: la ciencia cognitiva, los sistemas expertos, la inteligencia artificial, las redes neuronales, el trabajo en grupos apoyado por computadoras, la ciencia de la información, las redes semánticas e incluye a los sistemas de soporte a las decisiones.

Nos encontramos en una etapa en la cual la tecnología provoca cambios en la estructura y funcionamiento de todas las organizaciones. No obstante lo difundido de su utilización para las tareas operativas, este tipo de herramientas permanece subutilizadas por los líderes de las organizaciones en los procesos de toma de decisión⁶.

2. Materiales y métodos

2.1. Minería de Datos

La Minería de Datos (*Data Mining*) es un proceso que, a través del descubrimiento y cuantificación de relaciones predictivas en los datos, permite transformar la información disponible en conocimiento útil para la organización. Dentro de los sistemas de soporte a las decisiones hallamos a la Minería de Datos como una herramienta poderosa. La minería de datos es una idea basada en una simple analogía. El crecimiento de los almacenes de datos ha creado montañas de datos, que representan un recurso de gran valor para la organización. No obstante para extraer lo valioso de la montaña se debe excavar, o practicar la minería, para llegar a las ‘pepitas’ de metal precios, en nuestro caso el conocimiento⁷.

En todas las organizaciones, aún en la Universidad como así también en los ministerios del gobierno, las bases de datos son construidas con gran entusiasmo, no obstante no existe consenso acerca del significado de la minería de datos o de lo que dicho proceso conlleva, o bien de cuál es el resultado del proceso, ni de qué herramientas se vale.

⁶ MC. NURLIN, B.; SPRAGUE, R. **Information systems management in practice**. 4. Ed. N. J.: Prentice Hall, 1998.

⁷ FAYYAD, U.; PIATETSKY-SHAPIRO, G; SMYTH, P.; UTHURUSAMY, R. **Advances in Knowledge Discovery and Data Mining**. Cambridge, MA: MIT Press, 1996.

La minería de datos, consistente en la extracción de información oculta y predecible de grandes bases de datos, es una poderosa tecnología con gran potencial para ayudar a las compañías a concentrarse en la información más importante de sus bases de información⁷.

Al Descubrimiento de Conocimiento en Bases de Datos (Knowledge Databases Discovery KDD) se lo define como al Proceso de extracción no trivial para identificar patrones que sean válidos, novedosos, potencialmente útiles y entendibles, a partir de datos⁸.

Las técnicas más comúnmente usadas en minería de datos son las redes neuronales artificiales, los árboles de decisión, algoritmos genéticos y la regla de inducción⁹.

2.2. Redes Neuronales Artificiales

Como herramientas de la Minería de Datos las Redes Neuronales Artificiales han ganado una creciente importancia. Algunas definiciones sobre estos algoritmos: para Simon Haykin¹⁰ "Una red neuronal es un procesamiento distribuido masivamente paralelo que tiene una tendencia natural para almacenar conocimiento empírico y hacerlo disponible para el uso. Recuerda al cerebro en dos aspectos: (1) Conocimiento se adquiere por la red a través de un proceso de aprendizaje; (2) Las conexiones interneurónicas se conocen como pesos sinápticos y se usan para almacenar el conocimiento".

Para poder aprender, las redes neuronales se sirven de un **algoritmo de aprendizaje**. Estos algoritmos están formados por un **conjunto de reglas** que permiten a la red neuronal **aprender** (a partir de los **datos** que se le suministran), mediante la modificación de los **pesos sinápticos** de las conexiones entre las neuronas (recordar que el umbral de cada neurona se modificará como si fuese un peso sináptico más).

⁸ PIATETSKY-SHAPIRO, G.; FRAWLEY, W. **Knowledge Discovery in Databases**. AAAI Press, The MIT Press, 1991.

⁹ FAYYAD, U.; PIATETSKY-SHAPIRO, G.; SMYTH, P. Knowledge discovery and data mining: towards a unifying framework. In: SIMOUDIS, E., HAN, J., AND FAYYAD, U., ed. **Proceedings of KDD'96: Second International Conference on Knowledge Discovery & Data Mining**. CA: AAAI Press, Menlo Park, 1996. p. 82-88.

¹⁰ HAYKIN, S. **Neural Networks**. Macmillan College Publishing Company, 1994.

Generalmente los **datos** que se usan para entrenar la red se le suministran de manera aleatoria y secuencial.

2.3. Redes Neuronales Auto-organizadas

Este grupo de redes neuronales se caracteriza porque en su entrenamiento no se presentan las salidas objetivo que se desean asociar a cada patrón de entrada. La red, a partir de un proceso de auto-organización, proporcionará cierto resultado, el cual será reflejo de las relaciones de similitud existentes entre dichos patrones de entrada. La principal aplicación de estos modelos, será la realización de agrupamiento de patrones, visualización de datos y representación de densidades de probabilidad.

A diferencia de lo que sucede en el aprendizaje supervisado, en el no supervisado (auto-organizado) no existe ningún maestro externo que indique si la red neuronal está operando correcta o incorrectamente, pues no se dispone de ninguna salida objetivo hacia la cual la red neuronal deba tender.

2.4 Las Redes autoorganizadas creadas por Kohonen

Teuvo Kohonen presentó en 1982 un modelo de red neuronal con capacidad para formar mapas de características de manera similar a como ocurre en el cerebro¹¹.

El objetivo de Kohonen era demostrar que en un estímulo externo (información de entrada) por si solo, suponiendo una estructura propia y una descripción funcional del comportamiento de la red, era suficiente para forzar la formación de mapas^{12, 13}.

Las redes presentadas por Kohonen pertenecen a la categoría de las redes competitivas o mapas de autoorganización, es decir, aprendizaje no supervisado. Poseen una arquitectura de dos capas (una de entrada y la segunda de salida), presenta funciones de activación lineales y flujo de información unidireccional

Las unidades de entrada reciben datos, vectores de patrones normalizados, se normalizan así mismo los pesos de las conexiones con la capa de salida. Tras el aprendizaje de la red, **cada patrón de entrada activará una única unidad de salida.**

¹¹ KOHONEN, T. The self organizing map. **IEEE Proceedings**, v. 78, n. 9, p. 1464-1480, 1990.

¹² KOHONEN, T. Physiological interpretation of the self-organizing map algorithm. **IEEE Trans. on Neural Networks** 6, 895-905, 1993.

¹³ KOHONEN, T. Self-organized formation of topologically correct feature maps. **Biological Cybernetics**, v. 43, p. 59-69, 1982.

El objetivo de este tipo de redes es clasificar los patrones de entrada en grupos de características similares, de manera que cada grupo activará siempre la(s) misma(s) salida(s). **Cada grupo de entradas queda representado en los pesos de las conexiones de la unidad de salida que activa.** La unidad de salida ganadora para cada grupo de entradas no se conoce previamente, es necesario averiguarlo después de entrenar a la red.

El mapa auto-organizativo ó SOM ha probado ser una herramienta valiosa en Minería de Datos y KDD en aplicaciones de análisis de datos financieros. Se ha aplicado en forma exitosa en problemas de ingeniería tales como reconocimiento de patrones, análisis de imágenes, monitoreo y diagnóstico de fallas¹⁴.

El SOM tiene ciertas características que lo hacen un método útil para la Minería de Datos. Lleva a cabo un mapeo que reduce la dimensionalidad en forma ordenada del conjunto de entrenamiento. El mapeo es coherente con la función densidad de probabilidad de los datos y presenta un comportamiento robusto con los datos perdidos. Es sencillo de explicar y, aún más importante, de visualización fácil.

A partir de experiencias se puede inferir el éxito de la aplicación de los mapas SOM otras áreas de la ciencia^{15, 16, 17, 18}.

3. Resultados

Con el fin de experimentar técnicas de minería de datos se propone la utilización de la bases de datos académicos de los alumnos de la carrera de Ingeniería Electrónica, inscriptos en el Plan 1992.

¹⁴ VESANTO, J. SOM- based data visualization. **Intelligent Data Analysis** 3, p. 111-126, 1999.

¹⁵ PASSONI, L. I. *et. al.* Clustering algorithms as classifiers of blood pressure recordings. **Computer Simulations in Biomedicine Computational Mechanics Publications**. Boston.1995.

¹⁶ JOUTSINIEMI, S.; KASKI, S.; LARSEN, T. Self-organizing map in recognition of topographic patterns of EEG spectra. **IEEE Transactions on Biomedical Engineering**, 42:1062-1068, 1995.

¹⁷ KASKI, S.; KOHONEN, T. Exploratory data analysis by the Self-organizing map: Structures of welfare and poverty in the world. In: REFENES, A. P. N. *et al.* **Neural Networks in Financial Engineering. Proceedings of the Third International Conference on Neural Networks in the Capital Markets**. Singapore: World Scientific, 1996. p. 498-507.

¹⁸ Kaski, S. *et. al.* Creating an order in digital libraries with self-organizing maps. In: **Proceedings of WCNN'96, World Congress on Neural Networks**. Mahwah, NJ: Lawrence Erlbaum, INNS Press,.Lawrence Erlbaum and INNS Press, 1996. p. 814-817.

Hipótesis del experimento: El descubrimiento de relaciones de interés y la generación de una estrategia de visualización a emplear en la gerencia académica a partir de la evaluación multivariable disponible en la base de datos de alumnos.

Con el fin de poder analizar desde el punto de vista curricular el comportamiento académico, se decidió agrupar las materias del plan de estudios por áreas temáticas que hacen a la estructura curricular del área básica y específica de la carrera.

Con el propósito de lograr una estructura de datos adecuada para el posterior procesamiento se genera una única tabla cuya estructura de datos contiene: un campo por área que contenga el promedio obtenido en las materias que la componen y un campo por área que contiene la lentificación relativa al período en el que se espera que el alumno haya aprobado la última materia del área.

A partir de la obtención de las variables numéricas, se procedió a la estandarización estadística de las mismas con el objetivo de que sus magnitudes particulares no generen sesgos en la estructura del mapa auto-organizativo.

Para la programación del algoritmo del mapa auto-organizativo se utilizó el entorno Matlab 5.3[®]. Se utilizaron las librerías de dominio público generadas por el Centro de Investigación en Redes Neuronales de la Universidad Tecnológica de Helsinki.

El conjunto de entrenamiento del mapa S.O.M. contiene todos los casos (alumnos) que lograron completar el ciclo básico de la carrera, con la información de todas las áreas que se completan al tercer año de carrera (4 áreas, 8 variables: promedio y lentificación relativa de cada una de ellas).

La realización del mapa tiene por finalidad detectar las áreas topográficas donde se ubican el grupo de alumnos que finalizaron su carrera en el período esperado y dónde se halla el grupo de los que se considera desgrane (aquellos que a cuatro años de haberse terminado el Plan 92 no han terminado su carrera). En esta etapa se busca la relacionar esta ubicación en el mapa con el comportamiento (calificaciones y lentificación relativa) en las áreas básicas de la carrera.

A continuación se muestran resultados gráficos del experimento. En la figura 1, como resultado del proceso de aprendizaje, se muestra una grilla cuadrada de 25 celdas por lado, etiquetada con el fin de identificar las regiones a las que pertenecen los

mapa de la derecha de la misma figura se observa el histograma que pertenece a los que se consideran desgranados de su correspondiente cohorte (tipo D).

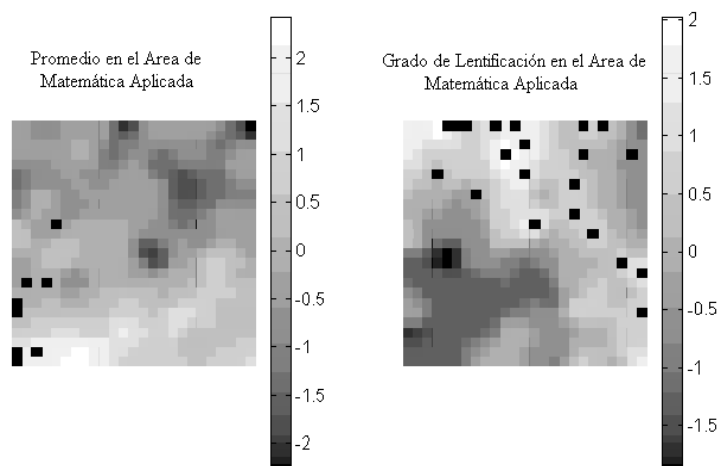


Figura 2

En el gráfico de la derecha de la figura 2 se observa la zona de impacto que pertenece a alumnos cuya tasa de lentificación es mucho mayor y en general los promedios son menores.

La interpretación del mapeo se debe realizar comparando la imagen de los mapas de variables individuales. A partir de esta interpretación se logró obtener parte de la explicación correspondiente a la lentificación en la carrera. Se investigó el comportamiento del Área con mayor lentificación (en este caso Matemática Aplicada) y como resultado del análisis se identificó la temática que causaba el problema. En consecuencia se propone estudiar detalladamente los contenidos y metodología de las materias que componen el área problemática.

4. Conclusiones

Como resultado de la experiencia se ha logrado extraer conocimiento interesante. A partir de una base de datos del tipo puramente cuantitativa se han observado situaciones de interés para la toma de decisiones del área académica, como



es: el grado de avance relativo junto con la calificación de una cohorte de alumnos desagregada por áreas del conocimiento.

Es importante considerar la posibilidad de evaluar un mayor número de atributos por alumno (edad, sexo, tipo de colegio del que proviene, condición socioeconómica), situación que posibilitaría el descubrimiento de nuevas relaciones, ciertamente interesante.

El impacto de los resultados es amplio, no desestimando la influencia económica, pues no se debe olvidar la estrecha relación existente en la influencia de la lentificación o grado de repitencia en el cálculo de los costos del egresado.