

Tecnologías Inteligentes para predicción de la aptitud del alumno. Una visión desde la práctica docente

Zulma Cataldi^{1,2}, Nancy Figueroa³, Fernando Lage^{1,2}
liema@fi.uba.ar, nnfigueroa@sinectis.com.ar, flage@fi.uba.ar

1. Facultad de Ingeniería. Universidad de Buenos Aires. Paseo Colón 850. C1063ACV Cdad. Aut. de Bs. As. ARGENTINA
2. Facultad Regional Buenos Aires. Universidad Tecnológica Nacional. SeCyT. Medrano 951. C1179AAQ Cdad. Aut. de Bs. As. ARGENTINA
3. Facultad de Informática y Ciencias de la Comunicación. Universidad de Morón. Cabildo 152. Morón. ARGENTINA

Resumen

En esta comunicación se presenta una investigación donde la acción se enfoca en diagnosticar el aprendizaje que permita mejorar la calidad de la enseñanza en el aula universitaria. Mediante la experiencia que se ha realizado hasta el momento se puede concluir que una red neuronal puede predecir el desempeño de los alumnos. Dado que esta es una tecnología poco explorada en ámbitos sociales, se considera importante su investigación debido al impacto que puede producir la predicción de escenarios críticos y en consecuencia efectuar la toma de decisiones para su solución, en este caso de aprendizaje. Los resultados dan cuenta de la apertura de nuevas posibilidades de aplicación de los sistemas inteligentes en educación.

Palabras claves: predicción, desempeño del alumno, redes neuronales.

I. Introducción

El aula universitaria, es hoy un espacio de carácter complejo dado por las interrelaciones entre los componentes, cuyas funciones dentro del sistema no son independientes y donde, el profesor debe ser un conocedor de la disciplina que desarrolla, pero también, debe saber *qué es lo que sucede en el aula, cómo aprenden los alumnos, cómo se puede organizar para ello el espacio y el tiempo, qué estrategias de intervención pueden ser más oportunas en ese determinado contexto.*

Luego, la práctica docente presenta aristas bien diferenciadas ya que el profesor debe cumplir, entre otras, las funciones de: *especialista en diagnóstico y prescripción del aprendizaje; especialista en recursos de aprendizaje, facilitador del aprendizaje, y especialista en la convergencia interdisciplinar de saberes.*

La investigación, enfoca la acción de determinar la evolución en sus aprendizajes mediante, precisamente, el desempeño del alumno como base de una predicción que permita mejorar la calidad de la enseñanza en tanto que será posible construir *planes de enseñanza a medida*. Cada estudiante, dentro de un enfoque centrado en el alumno, sigue un camino para el aprendizaje de acuerdo a sus fortalezas y debilidades.

El carácter complejo del aula universitaria suscitó la necesidad que las herramientas de predicción, fueran capaces de aprender cómo opera la realidad para realizar dicha predicción. Es así que se pensó en utilizar tecnologías inteligentes como las redes neuronales con estructura del tipo retropropagación (Cánepa *et al.*, 2006; Salgueiro *et al.*, 2006).

II. Objetivos

Se plantearon como objetivos: 1. Predecir la aptitud académica de los alumnos a través de redes neuronales y 2. Evaluar el sistema utilizando datos estadísticos.

El trabajo se centra en la idea que los patrones de entrada para la red representan las características de los alumnos en cuanto a sus logros en el dominio cognitivo.

III. Descripción del trabajo

a) Las redes Neuronales como herramientas de predicción:

Las redes neuronales son modelos diseñados para cumplir una cierta tarea y que en su funcionamiento intentan reproducir el comportamiento del cerebro humano. Las redes nacen debido a que hasta ahora, los sistemas creados en base a esquemas secuenciales no han logrado el cometido de *“interpretar el mundo”* y por ello, el esfuerzo de copiar a quien hasta ahora lo hace mejor: el cerebro, con su capacidad de procesamiento paralelo.

Las características funcionales que las hacen aptas se basan en los hechos de que las redes neuronales: a) Son sistemas robustos y tolerantes a fallas, que no afectan su desempeño, b) Son adaptables a nuevos ambientes por aprendizaje, en un sistema implicaría la necesidad de programarlo, c) Son capaces de manejar información poco específica o inconsistente, d) Puede ejecutar varias tareas simultáneas en paralelo y d) Son pequeñas, compactas y consumen poca energía

Kohonen (2001) plantea que: *“una red neuronal artificial es un conjunto de redes interconectadas masivamente en paralelo de elementos simples, usualmente adaptativos, y con organización jerárquica, las cuales intentan interactuar con los objetos del mundo real del mismo modo que lo hace el sistema nervioso biológico”*. El elemento clave en el paradigma de las redes neuronales artificiales son las estructuras que provee el sistema de procesamiento de la información.

Las redes neuronales artificiales no ejecutan instrucciones, sino que responden en paralelo a las entradas que se les presentan y el resultado no se almacena en una posición de la memoria, éste es el estado de la red para el cual se logra equilibrio. El conocimiento de una red neuronal no se almacena en instrucciones, el poder de la red está en su topología y en los valores de las conexiones (llamados pesos) entre neuronas. Claro está que las redes neuronales no alcanzan la complejidad del cerebro, sin embargo hay dos aspectos similares entre las redes que componen este órgano y las artificiales; primero que los bloques de construcción de ambas redes son elementos simples altamente interconectados; segundo, las conexiones entre neuronas determinan la función de la red.

Las redes neuronales se caracterizan por su topología, es decir por la disposición de las neuronas que forman la red dando lugar a capas o agrupaciones de neuronas que están, en mayor o menor medida alejadas de la entrada y la salida de la red. En este sentido, los parámetros fundamentales de red son: el número de capas, el número de neuronas por capa, el grado de conectividad y el tipo de conexiones entre neuronas.

2) Redes backpropagation

Backpropagation es un algoritmo de propagación de errores hacia atrás; es una regla de aprendizaje que se puede aplicar en redes con más de dos capas de neuronas y una de sus características es la capacidad de representar el conocimiento en las capas ocultas, obteniendo así cualquier correspondencia entre entradas y salidas (Del Brío y Sanz Molina, 2001).

El funcionamiento de este tipo de red backpropagation se basa en el aprendizaje de un conjunto predefinido de pares de datos entrada-salida dados a través de un ejemplo,

empleando un ciclo propagación y adaptación mediante dos fases. Primero se aplica un patrón de entrada en la capa de entrada de la red, que luego se propaga hacia las capas superiores hasta generar una salida, se compara el resultado obtenido en cada neurona de salida con el valor deseado para esa neurona y se obtiene un error para dicha unidad. A continuación, estos errores se transmiten hacia atrás, hacia todas las neuronas de la capa intermedia que contribuyen a la salida, recibiendo el porcentaje de error aproximado a la participación de la neurona intermedia en la salida original. Este proceso se repite capa por capa hasta llegar a la entrada y hasta que cada neurona haya recibido un error que describa su aporte al error total. Los pesos de las conexiones entre cada par de neuronas en la red se reajustan, según el valor del error recibido, de manera de que el error total cometido para ese patrón disminuya (Del Brío y Sanz Molina, 2001; Ohlsson, 1996 y Widrow y Lehr, 1990).

3) Estructura y aprendizaje de la red *backpropagation*.

Este tipo de red presenta una capa de entrada y de salida con una cantidad desigual de neuronas. Cada neurona de una capa intermedia recibe entrada de todas las neuronas de la capa anterior y genera salida hacia todas las neuronas de la capa siguiente (salvo las de salida). No hay conexiones hacia atrás (feedback) ni laterales (autorrecurrentes).

4) Los pasos para aplicar el algoritmo de entrenamiento

Para entrenar la red, en primer lugar debe darse valores a los pesos W_i ; de entrada que son aleatorios. Posteriormente, se debe exponer la red a un patrón de entrada y se especifica la salida deseada y luego, a partir de la obtención de la entrada neta y su respectiva salida de cada capa interna se puede obtener la salida de la red. La salida actual de la red se compara mediante una función apropiada al modo de representación de los datos con la salida deseada y se obtiene así el error a la salida. Si este error fue lo suficientemente grande, se propaga hacia atrás. Así, el error que se produce en una neurona oculta es proporcional a la suma de los errores conocidos que se producen en las neuronas a las que está conectada la salida de ésta, multiplicados por el peso de la conexión. Como último paso, se ajustan los pesos, mediante un algoritmo recursivo, desde las neuronas de salida a las de entrada (Del Brío y Sanz Molina, 2001).

5) La evaluación como instrumento de toma de datos.

Dadas las características de la tecnología a implementar se entiende por *predicción de la aptitud* al resultado de los alumnos en sus exámenes finales. De manera entonces que se busca la predicción del desempeño en exámenes finales.

Se trabaja sobre la idea que los datos de entrada representan las características de los alumnos en cuanto a sus logros en el dominio cognoscitivo definido en términos de conductas. Actualmente la universidad provee de sistemas formales de medición de logros que son las evaluaciones. Como establece Avolio de Colls (1987) una evaluación de logros cognoscitivos aborda diferentes aspectos, por lo que éstos se constituirían en los componentes de los vectores de entrada para la red neuronal. Estos aspectos serán: *información, comprensión, aplicación, síntesis, pensamiento crítico y trabajo intelectual*. A continuación se explica cómo se entienden cada uno de estos aspectos y se muestran los ejercicios que se implementaron en una evaluación de alumnos de la carrera de Ciencias Económicas para la asignatura Análisis Matemático I.

- *Información*: En este caso se busca que el alumno recuerde la información que leyó, de esta manera debe poner en evidencia el manejo de la terminología correspondiente al área. En la Figura 2, se muestra el ejercicio planteado en la evaluación. Aquí se busca que el alumno además de recordar el concepto de

función continua sea capaz de expresarlo en lenguaje matemático; es decir: Una función es continua en un punto $x = a$ si:

a) $\exists f(a)$; b) $\exists \lim_{x \rightarrow a} f(x) = L$, donde $L \in \mathbb{R}$; c) $f(a) = L$ lo que en palabras significa que la función existe; que el límite de la función cuando la variable tiende al punto a es finito y que éste último coincide con el valor de la función.

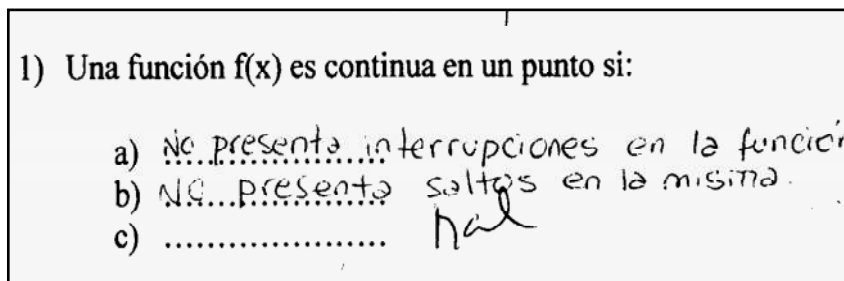


Figura 2. Ejercicio para evaluar la información de datos específicos

- *Comprensión*: Cuando un alumno comprende la información puede interpretarla y hacer uso de ella de tal manera que es capaz de traducir, interpretar, y extrapolar conceptos. En la Figura 3 se muestra el ejercicio correspondiente a este nivel del dominio cognoscitivo. En el mismo, se busca que el alumno interprete la información gráfica. De esta manera el debe traducir el lenguaje gráfico al matemático. En el primer dibujo debe expresar en lenguaje matemático que el límite funcional no es un valor finito. En el segundo dibujo debe expresar que si bien el límite funcional es un valor finito, éste no coincide con el valor de la función. Por último, en el tercer dibujo debe indicar que no existe el valor de la función en el punto indicado.

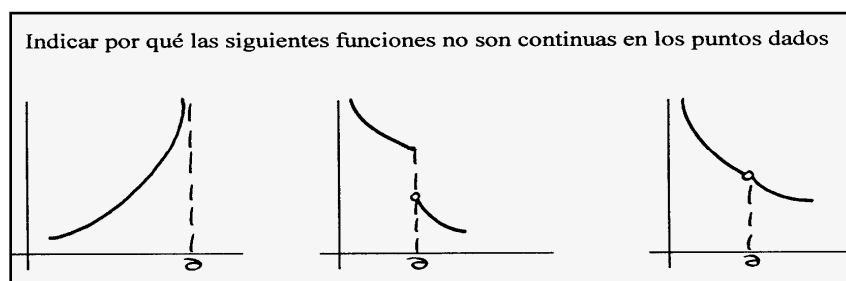


Figura 3. Ejercicio para evaluar la comprensión a través de la interpretación

- *Aplicación*: En este caso el alumno debe ser capaz de resolver situaciones no del todo conocidas, puesto que sino sería suficiente la memorización. En la Figura 4, se muestra el ejercicio correspondiente. El ejercicio propone al alumno una situación ficticia donde debe aplicar el concepto de límite de funciones, ya que el *tiempo* dado como variable no puede adoptar el valor infinito al que se refiere el problema cuando dice *extiende indefinidamente*. También debe armar dicha función.

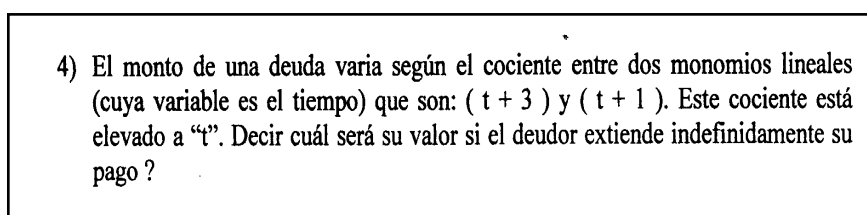


Figura 4. Ejercicio para evaluar la habilidad para aplicar

- *Síntesis*: Implica la elaboración creadora y personal de un producto reuniendo partes y elementos para formar un todo. En el ejercicio de la Figura 5 se busca que el alumno por sí mismo implemente una secuencia de pasos tendientes a verificar la igualdad propuesta.

15) Verificar que se cumple la siguiente igualdad:

$$\frac{d \left\{ \ln \sqrt{\frac{1+\sin x}{1-\sin x}} \right\}}{dx} = \sec x$$

Figura 5: Ejercicio para evaluar el nivel de síntesis

Como existen varios caminos para que el primer miembro de la igualdad llegue a transformarse en el segundo miembro de la igualdad, se espera que cada alumno ofrezca una alternativa diferente entre las posibles.

- *Pensamiento crítico*: Se evalúa si el alumno es capaz de juzgar el valor de una idea, obra, solución, método o material. En el caso de la Figura 6, cuya respuesta es correcta, se pide al alumno que evalúe varias respuestas para determinar cual es la correcta de acuerdo al criterio que es que la derivada primera de la función algún punto de los propuestos sea igual al valor de la función en el mismo; es decir que: $f'(x) = f(x)$. En este ejercicio también es necesario que el alumno comprenda la información gráfica planteada y además porque debe construir mentalmente rectas tangentes a la curva dada y establecer que la tangente trigonométrica de su ángulo es numéricamente igual al valor de la función del punto que él indique como correcto.

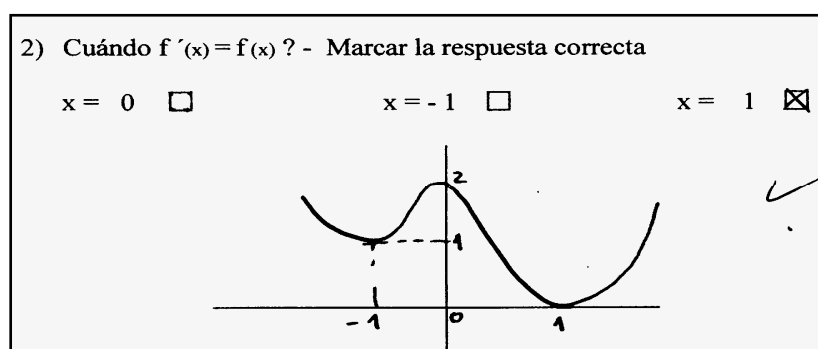


Figura 6. Ejercicio para evaluar el pensamiento crítico

- *Trabajo intelectual*: En este caso se espera que el alumno sea capaz de investigar, estudiar, leer, realizar lecturas críticas y obtener información de diferentes fuentes. Si bien esta habilidad puede considerarse como producto de las anteriores, es relevante que las utilice en pequeñas investigaciones. La Figura 7 muestra un caso aplicable en el aula como complemento a una evaluación escrita.

Investigue en diferentes fuentes como se deduce el Teorema de Rolle, Lagrange y Cauchy. Evalúe, en caso de existir, las diferencias que presentan las mismas.

Figura 7. Ejercicio para evaluar la capacidad para el trabajo intelectual

6) Descripción de la aplicación

Con este nuevo planteo de evaluación, se busca representar el estado cognitivo del alumno de la manera más apropiada (Figuroa *et al.*, 2007; Lage *et al.*; 2007). Luego, se deben adaptar los datos para que sean susceptibles de trabajarlos en la red neuronal ya que solo aceptan 0 y 1. Se seleccionó la aplicación NNClass que es una implementación del algoritmo Backpropagation (Saha, 1998). La misma permite trabajar con hasta cincuenta variables de entrada y es posible trabajar con variables tipo categóricas o continuas. En su implementación en Excel se pueden ver diferentes hojas para completar los datos que servirán para definir y construir la red (Hoja UserInput), ingresar los datos con los que el algoritmo definirá los pesos (Hoja Data). Además, presenta la hoja Calc, en la cual, se muestra el conjunto de pesos finales una vez construido el modelo. También mostrará los valores de los errores cuadráticos del entrenamiento y la validación en la hoja denominada Output. Se estableció en seis el número de entradas que corresponden a las variables *información, comprensión, aplicación, síntesis, pensamiento crítico y trabajo intelectual* que es la propuesta desde el marco pedagógico. El número de capas ocultas (*Hidden Layer*) se estableció en uno, con dos neuronas. Como la muestra disponible es de treinta y un exámenes, estoas son los *Total rows*. Los restantes ítems correspondientes a las *opciones de entrenamiento* adquieren los valores predeterminados (Figura 7).

Network Architecture Options

Number of Inputs (between 2 and 50): **6**

Number of Hidden Layers (1 or 2): **1**

Learning parameter (between 0 and 1): **0.9**

Momentum (between 0 and 1): **0.1**

Hidden Layer sizes (Maximum 20): **Hidden 1: 2, Hidden 2: 0**

Initial VM Range (0 +/- w): w = **0.7**

Training Options

Total #rows in your data (Minimum 10): **31**

No. of Training cycles (Maximum 500): **50**

Present inputs in Random order while Training?: **NO**

Training Mode (Batch or Sequential): **Sequential**

Saving Network Weights

With least Training Error

Training / Validation Set

Use whole data as training set

If you want to partition, how do you want to select the Validation set?

Please choose one option: **1**

Option 1: Randomly select 8% of data as Validation set (between 1% and 50%)

Option 2: Use last 8 rows of the data as validation set

Please fill up the input necessary for the selected option

Save model in a separate workbook?: NO

Build Model

Figura 7: Pantalla en la que se ingresan las características de la arquitectura de una red tipo NNClass

A partir de los registros obtenidos, se destina la tercera parte al entrenamiento. La idea es que una vez que la red aprende se ingresa un nuevo conjunto de valores, cuya salida es conocida pues pertenece a la dos terceras partes de las observaciones, la predicción de la red entonces puede ser evaluada.

7) El aporte de los sistemas de predicción a la práctica docente

ID	Species Name	Cont	Cont	Cont	Cont	Cont	Cont
1	aprobado	111	111	111	111	111	111
2	aprobado	100	111	10	11	1	111
3	aprobado	111	111	10	1	11	110
4	aprobado	1	111	100	101	11	110
5	desaprobado	0	0	11	11	0	11
6	aprobado	111	111	110	110	111	111
7	aprobado	111	1	1	0	0	110
8	desaprobado	1	0	1	0	0	1
9	desaprobado	0	1	0	1	0	0
10	aprobado	111	110	1	10	111	111
11	aprobado	111	111	111	1	100	100
12	aprobado	11	110	0	111	111	110
13	desaprobado	100	100	111	110	0	111
14	aprobado	111	111	111	110	10	11
15	aprobado	101	101	111	100	111	100
16	aprobado	111	111	100	100	101	111
17	aprobado	111	101	111	101	100	1
18	aprobado	100	110	101	101	110	100
19	desaprobado	100	110	0	1	0	0
20	desaprobado	10	10	11	100	11	110
21	desaprobado	0	100	0	100	1	0
22	aprobado	111	111	11	11	1	110
23	aprobado	10	1	1	1	0	11
24	aprobado	111	11	10	11	1	101
25	aprobado	111	101	1	1	1	0
26	desaprobado	110	110	1	0	0	0
27	desaprobado	101	11	0	0	0	10
28	aprobado	111	11	11	1	10	11
29	aprobado	111	111	110	11	10	11
30	aprobado	111	110	11	11	10	11
31	desaprobado	11	10	10	11	0	0

Figura 8: Datos de ingreso

La última etapa consiste en la obtención de los resultados del proceso a partir de la carga de los datos iniciales. En la Figura 8 se muestra la pantalla de con datos de entrada donde se observan las columnas desde E1 hasta E6 que son los datos de ingreso (input). En la columna nombrada como output se indica la salida esperada: aprobado, desaprobado.

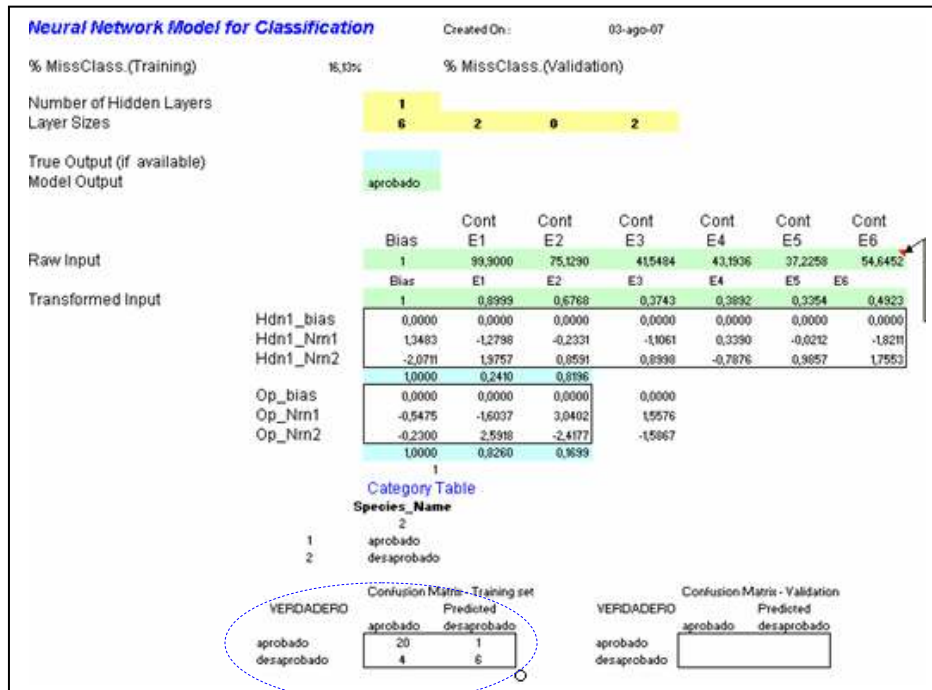


Figura 9: La salida de la red

En la Figura 9 se ve la probabilidad de acierto de la red para el juego de pesos y la muestra analizada está en el orden del 83% (20/31+6/31). La predicción de desempeño de los alumnos ayuda a la tarea docente debido a que, se reduce la elección de las estrategias de enseñanza y aprendizaje específicamente a las necesidades del alumno; además la planificación que el docente realiza también será más ajustada ya que no habrá dilaciones en el tiempo o cambio sustanciales de estrategias. Por otro lado, el hecho que el profesor conozca las necesidades de su grupo con suficiente antelación hará que mejore el clima del aula pues el alumno no estará sujeto a las tensiones propias de aquellos que tienen inconvenientes en el aprendizaje. En síntesis, la enseñanza y el aprendizaje resultarán más efectivos.

IV. Conclusiones

1. Se determinó que la aplicación NNclass era la más apropiada con un algoritmo de tipo retropropagación y esta decisión se justifica a través del análisis que surge del estado de las investigaciones donde resulta ser la más probada para la predicción.
2. Se creó la base de datos con los resultados del primer parcial, segundo parcial y los exámenes finales. Éste examen fue diseñado para evaluar seis dimensiones Información (comprensión, aplicación, síntesis, pensamiento crítico y trabajo intelectual) del logro cognitivo de los alumnos como indicadores de su estado de conocimiento; variable desglosada sobre la cual se realiza la predicción. Se tomó como grupo de observación a los alumnos de un curso de Análisis Matemático I de

Ciencias Económicas, siendo una muestra no probabilística de 31 sujetos disponibles. En este punto la dificultad estuvo en que los alumnos no fueron expuestos a modalidades de evaluación como la propuesta, de manera que se hizo necesario un reajuste para el segundo parcial, tratando de respetar los modelos tradicionales.

Los inconvenientes encontrados en el desarrollo de esta investigación han sido de diferente índole por lo que se cree pertinente su enumeración:

- El relevamiento de los tipos de exámenes se tornó un tanto dificultoso debido a que no existen procedimientos formales en la universidad para su construcción y archivo, por lo que se optó un camino inverso, es decir proponer evaluaciones de acuerdo a las variables que arroja el marco teórico como factores del estado cognitivo del alumno. De este modo se produjo la toma de datos para el primer parcial. En este punto se presentó como dificultad que los alumnos nunca fueron expuestos a modalidades de evaluación como la propuesta, por lo que fue necesario un reajuste para el segundo parcial, tratando de respetar los modelos tradicionales. Si bien se tuvo cuidado que los instrumentos fueran equivalentes, es de esperar que el error en la predicción de la red se pueda modificar por este motivo.
- Debido a que la normativa de la Universidad en la que se tomaron los datos, permite a los alumnos rendir sus exámenes finales durante un lapso de tiempo de más de un año lectivo, la estimación de la predicción con los datos de la totalidad de la muestra, se extenderá hasta esa fecha, superando el tiempo estimado.

Queda como línea a seguir, llevar a cabo la base datos con ejercicios para recomendar al estudiante y sugerir al docente las estrategias que mejor ajustan a las necesidades de sus alumnos.

V. Agradecimientos

Esta comunicación es parte de los proyectos de investigación: *Sistemas inteligentes aplicados a la predicción del comportamiento de los estudiantes y diagnóstico* 2005-2006 LIE-DC/04-07 del Laboratorio de Informática Educativa y Medios Audiovisuales (LIEMA) de la Facultad de Ingeniería, de la Universidad de Buenos Aires en transferencia a la Universidad de Morón, Facultad de Informática y Ciencias de la Comunicación, proyecto *Redes neuronales para predecir el alumno y sugerir acciones*. 2006-2007 Código B01-006/06 y el Proyecto *Modelado del tutor basado en redes neuronales para un Sistema Tutor Inteligente* Código 25/C099 del Programa de Incentivos 2007-2008 en el marco del Convenio FI-UBA y UTN-FRBA.

VI. Referencias

- Avolio de Colls, S. (1987) *Evaluación del proceso de enseñanza-aprendizaje*. Ediciones Marymar. Bs. As.
- Cánepa, S., Cataldi, Z.; Salgueiro, F., Lage, F. y García-Martínez, R. 2006. *Neuronal networks applied to the prediction of the behaviour of the students and diagnosis*. Aceptado WCCSETE'2006: World Congress on Computer Science, Engineering and Technology Education. Santos. Brasil. Marzo 19-22.
- Del Brio; B. y Sanz Molina, A. (2001): *Redes neuronales y sistemas difusos*. Paraninfo.

- Figuerola, N.; Cataldi, Z. y Lage, F. 2007. *La tecnología informática como aporte interdisciplinario al mejoramiento de las herramientas de la evaluación*. II Jornadas Nacionales y I Latinoamericanas de Pedagogía Universitaria "Enseñar y aprender en la Universidad. Culturas y educación en la Universidad: problemas y perspectivas". 6 y 7 de setiembre. Universidad Nacional de San Martín.
- Kohonen, T. (2001). *Self-Organizing Maps, third edition*. Springer series in information sciences. Ed Springer. Helsinki University of Technology Neural Networks Research Centre P.O. Box 5400 02
- Lage, F., Cataldi, Z. y Figuerola, N. 2007. *Redes neuronales para predecir la aptitud del alumno y sugerir acciones*. ICECE 2007: International Conference on Engineering and Computer Education. Monguaguá and Santos, Brazil. March 11-14. Book of Abstracts published by Claudio da Rocha Brito (ISBN 85-89120-43-0) & Melany M. Ciampi (ISBN 85-89549-39-9).
- Ohlsson, S. (1996) Learning from performance of errors. *Psychological Review* 3 (2) p. 241-262.
- Saha, A. (1998) *Application of Ridge Regression for Improved Estimation of Parameters in Compartmental Models*; Tesis Doctoral. Departamento de Estadística;
- Salgueiro, F.; Costa, G; Cánepa, S.; Lage, F; Kraus, G; Figuerola, N. y Cataldi, Z. 2006. *Redes Neuronales para predecir la aptitud del alumno y sugerir acciones*. Anales del VIII Workshop de Investigadores en Ciencias de la Computación. Páginas 645-648. ISBN 13: 978-950-9474-34-5, 1 y 2 de junio.
- Widrow B., Lehr M. *30 years of adaptive neural networks: Perceptron, Madaline. and Backpropagation*. Proceedings of the IEEE, vol 78 Num. 9 Septiembre de 1990. pp 1415-1442.