

REDES NEURONALES ARTIFICIALES APLICADAS A LA EVALUACIÓN DOCENTE Y A LA TOMA DE DECISIONES EN MATEMÁTICA EDUCATIVA

Martínez Luaces, V. Martínez Luaces, F
Universidad de la República. Universidad ORT. Uruguay.
victorml@fing.edu.uy , f.martnez@mailcity.com

Introducción y antecedentes

La evaluación ha sido un tema de constante preocupación en las diferentes áreas educativas, pero particularmente en las Matemáticas. Mucho se ha escrito sobre la evaluación del rendimiento estudiantil con relación a esta área del conocimiento, pero son considerablemente menos frecuentes los trabajos que se ocupan de la evaluación del desempeño de los docentes.

En nuestro caso, ésta ha sido una línea de investigación que tiene ya varios años de desarrollo y en la que aún continuamos. En diversos trabajos sobre el tema, hemos ido modificando la metodología para el tratamiento de los datos obtenidos.

En una primera instancia, en la Facultad de Ingeniería de la Universidad de la República de Uruguay, se utilizó un formulario de 25 preguntas con 5 opciones distintas de respuesta para cada pregunta, tomado de la Universidad de Valencia, España. Con este formulario, hemos recabado datos entre los años 1993 y 1996 inclusive. Con un instrumento esencialmente similar, pero esta vez utilizado en la Facultad de Química de la misma Universidad, se obtuvieron datos entre los años 1995 y 1997 inclusive. En los hechos, el cuestionario fue modificado para atender a la diversidad de situaciones: cursos teóricos, cursos prácticos de ejercicios y cursos prácticos de laboratorio.

En base a los datos obtenidos en la Facultad de Química, se realizó un primer trabajo, con estadísticas elementales, básicamente descriptivo, en esta área de investigación (Martínez Luaces, V., 1998a). Anteriormente se trabajó con un grupo de expertos, obteniéndose conclusiones muy similares a las que resultaron, posteriormente, de la opinión estudiantil (Casella, S. y Martínez Luaces, V., 1996)

De todos estos insumos surgió un trabajo de carácter más general, presentado en el Grupo de Estudio sobre la Enseñanza y el Aprendizaje de la Matemática en el Nivel Universitario, organizado por ICMI en Singapur a fines de 1998 (Martínez Luaces, V., 1998b).

Años más tarde, se volvió a trabajar sobre Evaluación Docente (Gómez, A. y Martínez Luaces, V., 2001) y Evaluación de la Calidad de Enseñanza (Gómez, A., Guineo, G. y Martínez Luaces, V., 2002), en la Facultad de Química, utilizando para el tratamiento de los datos, técnicas de la Estadística No Paramétrica y del Análisis Multivariado.

Metodología

A principios del año 2002, profundizamos en un nuevo aspecto de esta área de investigación, al utilizar una metodología diferente: el análisis de la evaluación docente mediante el uso de Redes Neuronales Artificiales (ANN por sus siglas en

inglés). En este trabajo se analizan datos de Facultad de Ingeniería y de Facultad de Química, intentando predecir el juicio global de los estudiantes a partir de los resultados de otras variables.

¿Por qué la idea de utilizar ANN en relación a la Evaluación Docente? Antes de responder a esta pregunta comenzaremos por describir brevemente en qué consisten y cuál fue su origen. Las ANN, como parte de la llamada computación no-lineal, originadas hace algunas décadas, representan un intento de copiar los procesos (con alto grado de paralelismo) que tienen lugar en las neuronas del cerebro, de acuerdo a investigaciones procedentes de la Neurobiología. Esta herramienta ha tenido gran difusión en los últimos años, debido al éxito que se ha obtenido con ésta técnica en diversas áreas, al punto que en muchas de ellas, las ANN ya han abandonado el laboratorio, para implementarse en forma industrial. Son muy diversos los problemas para los cuales no se encontraban algoritmos lineales satisfactorios, y sin embargo, a través de las ANN se ha arribado en muchos de ellos a una solución factible, en particular cuando dichos problemas tienen que ver con el reconocimiento de patrones. En el estudio de la Evaluación Docente cabría preguntarse si existen patrones que permitan medir la eficacia docente. Al respecto, es común que las encuestas realizadas entre el estudiantado soliciten una evaluación parcial del docente en diversos aspectos de su desempeño. Algunos de estos aspectos son: conocimiento disciplinar, manejo del pizarrón, motivación, relación con los estudiantes; etc.. A todo lo anterior se agrega una evaluación global del docente.

Una situación ideal sería, sin duda, determinar si existe una relación funcional entre dichas variables parciales y el juicio global. En la práctica esto no es viable, no solamente por la complejidad matemática que tendría dicha función, sino porque ésta no es estática en tiempo y espacio, sino que en diferentes dominios (en nuestro caso encuestas en diferentes instituciones) y también en diferentes períodos, dicha relación funcional experimentará cambios.

Minimización del error en Perceptrón Multicapa

El Perceptrón es un tipo de ANN que, como otros, se compone de unidades de procesamiento llamadas “neuronas” o elementos de proceso (PE, en inglés). Cada neurona tiene en general varias entradas y una única salida de datos, aunque esta puede distribuirse entre varias neuronas. A su vez, dichas entradas se ven afectadas por un vector de “pesos” como puede verse en la figura 1.

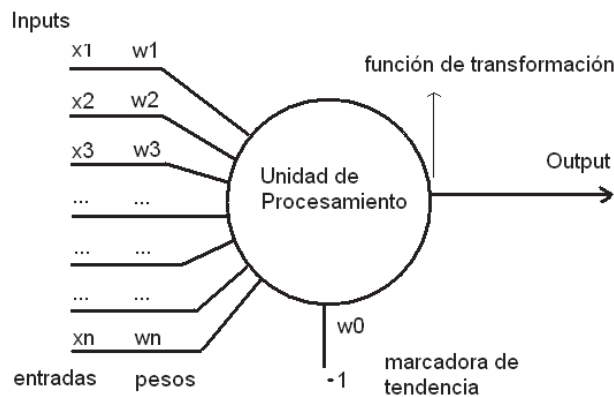


Figura 1.

Las neuronas se organizan en estructuras de datos llamadas “capas”. Cada capa recibe información de la capa anterior y la pasa a la siguiente. La capa inicial o Capa de Entrada, es la que toma, en nuestro caso los datos de la encuesta, luego se interconecta con la capa siguiente llamada Capa Oculta, que es donde ocurre la parte fundamental del proceso y finalmente hay una Capa de Salida, donde en nuestro caso, se obtiene el juicio global resultante. Todo lo anterior se encuentra esquematizado en la figura 2.

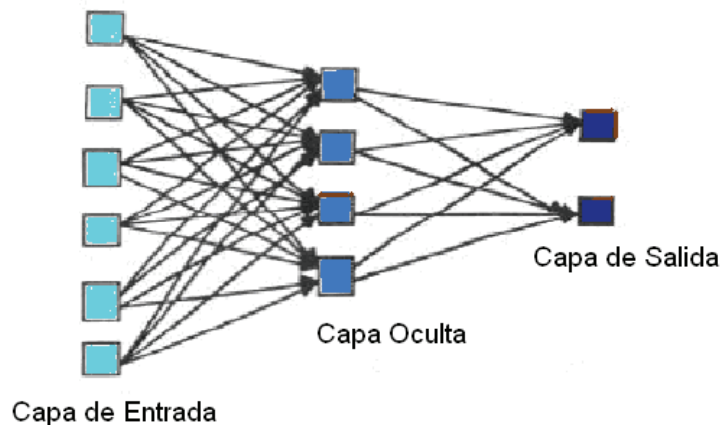


Figura 2.

La clave de la capacidad de las ANN para predecir resultados, radica en los pesos (valores numéricos) que en cada neurona se asignan a las variables que se le introducen. A su vez, dentro de cada neurona, se realiza una suma ponderada de sus entradas con dichos pesos, y luego de transformar el resultado mediante una función específica se pasa a las neuronas siguientes. En nuestra ANN, fijamos estos pesos inicialmente mediante una función randómica en valores pequeños, pero luego, a medida que se “entrena” la red, éstos deben ir variando hasta alcanzar una configuración óptima que minimice el error del resultado final (Picton, Ph, 2001).

El proceso anteriormente descrito, se logra mediante el uso de un algoritmo llamado “BackPropagation” o Propagación hacia atrás, que es el que utilizamos. Dicho algoritmo consiste en entrenar la red, dándole juegos de variables, y luego de obtener un resultado global por parte de la ANN, indicarle cual debió ser el resultado exacto, para que realice correcciones tendientes a minimizar la función de error. En la neurona de salida, esto es inmediato, ya que al calcular la diferencia, pueden ajustarse sus “pesos” para que en una nueva ejecución, el error sea menor. Luego, en las neuronas de capa oculta, el error no es explícito, porque no podemos determinar directamente cual debió ser la salida de cada una para que el resultado final fuese exacto. Aquí es donde se aplica el algoritmo de BackPropagation, tomado de la Teoría de Errores, en particular aplicando el cálculo del error de las variables en conexión con un error funcional dado. En las ANN elementales, se suele utilizar la función “escalón” como función de transformación de la salida, pero en nuestro caso, utilizamos la función “sigmoïdal”, tangente hiperbólica, ya que tiene un

comportamiento que se aproxima a dicha función pero con la ventaja respecto a aquella de ser derivable, un requisito esencial para implementar la Propagación Hacia Atrás. Así, a medida que se entrena la red, se busca minimizar el error en cada neurona, lo que significa minimizar las derivadas, por lo que también se le ha llamado a esta técnica “algoritmo del gradiente descendente”.

Resultados

En este trabajo, se utilizaron 12 variables parciales y un juicio global obteniéndose promedios de error del 2% y menores. En la figura 3 se puede apreciar la evolución del error relativo con respecto al número de docentes considerados. Dicho error fluctúa al comienzo, luego disminuye y finalmente tiende a permanecer constante.

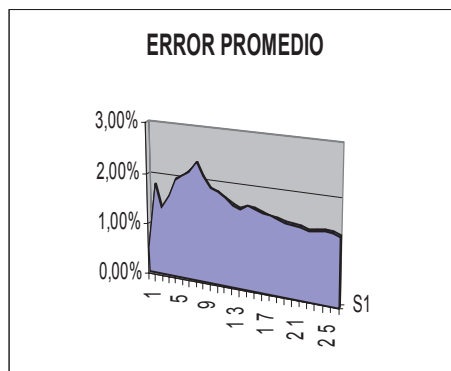


Figura 3. Error promedio vs. Número de docentes.

Además del error promedio, ya analizado, conviene visualizar el error cometido en cada valor predicho, respecto del valor real. En la figura 4 se representan los valores reales de juicio global para los 27 docentes considerados y el error absoluto correspondiente a cada caso.

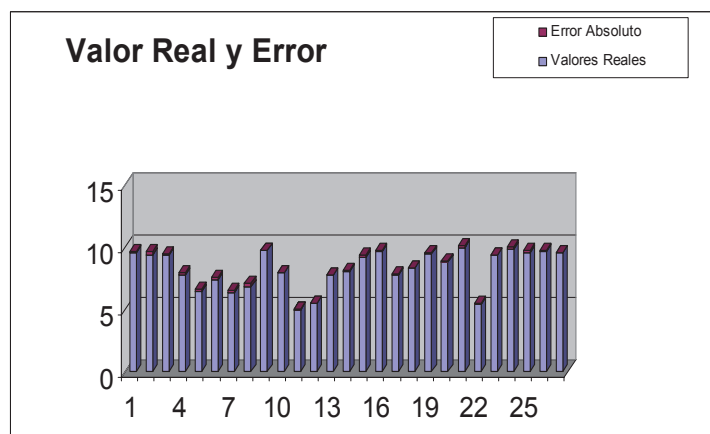


Figura 4. Valores reales de juicio global y error absoluto correspondiente.

Si bien los errores pueden ser considerados aceptables, un análisis estadístico mostró que los resultados predichos por la red y los valores reales no coinciden. En efecto, la aplicación del Test “t” de Muestras Ligadas, permite concluir que se trata de dos poblaciones diferentes cuando se trabaja al nivel de significación del 0.05 (Martínez Luaces, V., 1999). En una red que trabaja con 12 variables, como en este caso, la información introducida al sistema podría ser “excesiva”. Concretamente, podría contener información redundante que tiende a sobrevalorar alguna faceta por su repetición en diversas formas en el cuestionario. También, podría estar introduciendo “ruido” en la red debido a la inclusión de información no relevante que tiende a crear confusión, más que permitir definir los elementos que determinan la evaluación global.

Conclusiones

Del tratamiento de datos realizado surgen dos conclusiones inmediatas:

El error relativo tiende a estabilizarse entre 15 y 20 docentes considerados.

Los docentes con pocos alumnos encuestados (menos de diez) generan un mayor error en la aproximación

En el caso en estudio (12 inputs), se encontraron dificultades debido al número considerable de entradas, que paradójicamente, en el caso de ANN (y tal vez en la mente del estudiante que participa en la encuesta) provoca confusión al introducir información redundante o no relevante que puede ocultar los aspectos verdaderamente esenciales del desempeño docente. En el otro extremo, disponer de un número muy pequeño de preguntas no permitiría predecir adecuadamente un juicio global. Cabe entonces plantearse, la posible conveniencia de una solución de compromiso, donde se consideren los diversos aspectos de la práctica docente, pero evitando sobrevalorar alguno de ellos, o bien incluir información no relevante. Una posibilidad sería entonces, trabajar con “variables condensadas”, elaboradas en base a una “clusterización” previa (Gómez, A. y Martínez Luaces, V., 2001) y así poder, en principio, tomar como base una cantidad intermedia de variables fundamentales. Esta combinación de técnicas de Análisis Multivariado y ANN, justifica una cierta línea posible de continuación en este tipo de investigaciones.

Bibliografía

- Gómez A., Guineo, G. y Martínez Luaces, V., 2002 “Un trabajo de Investigación-Acción en un curso de Ecuaciones Diferenciales para Ingeniería de Alimentos”. *Página Web del X EMCI (Enseñanza de Matemática en Carreras de Ingeniería)*, Resistencia, Argentina.
- Gómez, A. y Martínez Luaces V., 2002. “Evaluación docente utilizando Análisis Multivariado”, *Acta Latinoamericana de Matemática Educativa*, 15.2, pp. 1016-1021. Ed. CLAME, México.
- Martínez Luaces, V. 1998a. Matemática como Asignatura de Servicio: algunas conclusiones basadas en una evaluación docente. *Números*. Revista Española de Didáctica de Matemática, España
- Martínez Luaces, V. 1998b. Considerations about Teachers for Mathematics as a Service Subject at the University *Pre-proceedings of ICMI Study Conference*, Singapore, Nanyang Technological University
- Martínez Luaces, V. y Casella, S., 1996 La educación matemática en las diferentes ramas de la Ingeniería en el Uruguay hoy. *Memorias del II Taller sobre la Enseñanza de la Matemática para Ingeniería y Arquitectura*. La Habana, Cuba.
- Martínez Luaces, V., 1999. *Estadística Aplicada a Ingeniería Ambiental*. Ed. IMFIA, Montevideo.
- Picton, Ph., 2001. *Neural Networks*, pp. 37-47.