

USO DE UNA RED NEURONAL ARTIFICIAL PARA LA PREDICCIÓN DEL RENDIMIENTO ACADÉMICO: UN ESTUDIO PRELIMINAR.

JOAQUÍN DOMÍNGUEZ
HUMBERTO ACEVEDO
UNIVERSIDAD CENTRAL
DE VENEZUELA

RESUMEN

En este trabajo se desarrolló una red neuronal artificial para predecir, con un alto grado de exactitud, el rendimiento académico. En la implementación de la red se utilizó una arquitectura de tres capas con ocho unidades de entrada, diezésis (16) unidades intermedias y una unidad de salida. La aplicación del modelo predictivo se ilustra con los resultados del p-ronóstico individual del rendimiento académico de los estudiantes del primer-año, curso 1995-1996, de la Escuela de Medicina "Luis Razetti" de la Universidad Central de Venezuela. Los datos, obtenidos retrospectivamente en una muestra de 220 estudiantes, fueron usados para "entrenar" y examinar la red. La red fue "entrenada" con un algoritmo de retropropagación para entrenamiento supervisado, con los datos correspondientes a la mitad de los estudiantes y fue examinada con los datos de la otra mitad, es decir, con datos que no intervinieron en la construcción de la red. Las predicciones fueron comparadas con los resultados reales obtenidos por los estudiantes en este año. Este cotejo se llevó a cabo con un análisis ROC. La red se desempeñó con una sensibilidad de 92% y una especificidad de 96%. El análisis computacional neuronal no lineal parece tener un gran potencial en la predicción del rendimiento académico, sin embargo, esta posibilidad debe ser confirmada con una muestra mayor de estudiantes.

Descriptores: Red neuronal - Rendimiento académico - Estudiantes de medicina - Análisis ROC.

USE OF AN ARTIFICIAL NEURAL NETWORK FOR THE PREDICTION OF ACADEMIC PERFORMANCE: A PRELIMINARY STUDY.

ABSTRACT

In this study an artificial neural network was developed to predict academic performance, with a high degree of accuracy. The network was implemented as a three-layer architecture with 8 input units, 16 internal or hidden units, and an output unit. The application of the prediction procedure is illustrated with the results of the individual prognoses of academic performance of the students from the first year, 1995-1996 course, at the "Luis Razetti" School of Medicine, Universidad Central de Venezuela. Retrospective data for a sample of 220 students to train and test the network were used. The network was trained, using a back propagation supervised training algorithm, on one half of the students and was tested on the remaining half of the students to which it had never been exposed. The predictions were compared with the real results obtained by the students in this year. For this the ROC analysis was used. The network performed with a sensitivity of 92% and specificity of 96%. The application of non linear neural computational analysis to the prediction of academic performance appear to have significant potential; however, this possibility must be confirmed through testing on a larger students sample.

Key words: Neural network-Academic performance-Medical students-ROC analysis.

INTRODUCCIÓN

En los actuales momentos existe una preocupación generalizada en torno al bajo rendimiento de los estudiantes en la educación superior. A pesar de la naturaleza selectiva del procedimiento de admisión implementado por las instituciones de este nivel educativo, casi invariablymente algunos miembros de cada cohorte de estudiantes fracasan en el primer año de la carrera universitaria. En este contexto, adquiere particular importancia la predicción del rendimiento académico de los estudiantes admitidos. Si se dispone de esta información, pueden tomarse medidas preventivas con los estudiantes en riesgo, en una etapa en la que aún es posible hacer previsiones que ayuden a conseguir un mayor grado de eficiencia en la labor docente.

La mayoría de los métodos estadísticos para la predicción del rendimiento académico encuentran problemas cuando los datos son extremadamente variables (Bacalao, 1985). Como una alternativa, en este trabajo, se describe una red neuronal artificial que ha sido "entrenada" para predecir el rendimiento académico de cada estudiante a partir de los datos que de él se poseen en el momento de su ingreso en la universidad.

En los últimos años, las redes neuronales artificiales han sido implementadas en un gran número de campos de la actividad humana (Schemajik y Triames, 1992; Yee y col., 1993; Baxi y Skora, 1996; Reichgelt, 1996; Stevens y col., 1996; Khan y col., 1998), demostrando una buena velocidad de procesamiento y una capacidad de generalización que las hace particularmente útiles para estudios predictivos (Canarrell, 1995). Ellas son "entrenadas" con las variables de entrada y, desarrollando los algoritmos necesarios en su arquitectura interna, producen un modelo predictivo sobre la base de los hechos aprendidos.

MATERIALES Y METODOS

MUESTRA

La experiencia consistió en tomar una muestra tomada por 220 estudiantes que ingresaron en la Escuela de Medicina «Luis Ruizelli» de la Universidad Central de Venezuela, en el año lectivo 1995-1996. La información porencialmente útil y disponible, por estar registrada en los expedientes en el momento de ingreso, se resume en las siguientes ocho variables: sexo, lugar de procedencia (Caracas o resto del país), institución de procedencia (oficial o privada), nivel socioeconómico, tiempo de graduado como bachiller promedio de las notas obtenidas en el bachillerato, puntaje obtenido en la prueba de razonamiento verbal y puntaje obtenido en la prueba de habilidad numérica. La información disponible de cada estudiante al finalizar el primer año de la carrera consistió en las notas finales (rango de 0 a 20) en las cuatro asignaturas siguientes: Anatomía, Bioquímica, Fisiología y Salud Pública.

A los efectos de considerar un indicador de lo que pudiera llamarse con propiedad "éxito académico" se creó una variable indicativa del resultado global (RG) en el primer año de la carrera, se realizó la siguiente definición: $RG = 1$, si el estudiante aprobó las cuatro asignaturas del primer año, $RG = 0$, si el estudiante reprobó una o más asignaturas del primer año.

RED NEURONAL ARTIFICIAL.

Una red neuronal artificial es un grupo de ecuaciones matemáticas interconectadas que aceptan datos de entrada (input) y que calculan un valor de salida (output) a partir de los datos de entrada. Esta estructura es un método pre-diseñado para la aplicación de técnicas estadísticas no lineales. La red está constituida por un cierto número de unidades individuales de procesamiento o neuronas. Cada neurona recibe una señal de entrada que procesa localmente a través de una función de activación y produce una señal de salida. Sin embargo, las neuronas de entrada no se conectan direc-

tamente con las neuronas de salida, sino a través de una o más capas intermedias de neuronas. Hay múltiples conexiones entre las neuronas de cada capa y entre las capas. Estas conexiones tienen fuerzas o "pesos" que son "aprendidos" por la red. El comportamiento de la red está determinado por los "pesos" de las interconexiones neuronales que son establecidos durante el proceso de entrenamiento de la red.

A través de un proceso experimental, la red puede ser "entrenada" para identificar patrones basados en un grupo de variables de entrada y producir un resultado de salida. Mas que ser programada, la red es "entrenada" presentándole grupos de datos junto con los resultados que el investigador desea asociar con los datos de entrada. La "capacidad de aprendizaje" y procesamiento de una red neuronal está determinada por la arquitectura de las interconexiones neuronales y el algoritmo de entrenamiento utilizado. Después de que ocurre el "aprendizaje", la red puede clasificar nuevos casos con las mismas o diferentes características. Estas particularidades hacen a las redes neuronales artificiales teóricamente capaces de "aprender" a partir de ejemplos y luego generalizar con casos que no han sido presentados antes (Wasserman, 1990).

CONSTRUCCIÓN DE LA RED NEURONAL ARTIFICIAL.

En la implementación de la red neuronal artificial se utilizó una red de tipo "feedforward", entrenada con el algoritmo de retropropagación y diseñada para activar su neurona de salida. La capa de entrada consistió en ocho unidades que representan las ocho variables que se conocían de cada estudiante en el momento de su ingreso. La capa intermedia tenía 16 unidades (no es posible producir a priori el número de unidades intermedias que sería requeridas para un funcionamiento óptimo), el número de unidades intermedias varió entre 2 y 20. La capa de salida consistió en una unidad única que representó el resultado de la predicción del rendimiento académico (éxito o fracaso en el primer año): 0,0 para fracaso y 1,0 para éxito.

La figura 1 ilustra la red usada en este estudio. Cada variable de entrada es representada por una neurona. Cada una de estas neuronas, a su vez, co-

necha con todas las neuronas de la capa intermedia. La fuerza o "pesos" de las relaciones entre las neuronas de las capas de entrada e intermedia es aleatoria al inicio del entrenamiento, pero en la medida en que la red obtiene experiencia con los datos, los "pesos" de las conexiones entre las neuronas cambian. Las variables normalizadas en las neuronas de la capa de entrada son multiplicadas por los valores de los "pesos" de conexión entre las neuronas de entrada y las intermedias. Si la suma de las señales de entrada a una determinada neurona intermedia alcanza un nivel umbral, se produce una señal de salida en la neurona intermedia que se relaciona con una función de transferencia que usualmente es no lineal. La señal de salida de la neurona intermedia es transmitida a la neurona de salida, nuevamente como una función de transferencia y el respectivo "peso" de conexión.

ENTRENAMIENTO DE LA RED NEURONAL ARTIFICIAL

El proceso de entrenamiento de la red consistió en la selección retrospectiva de los datos correspondientes a la mitad de los estudiantes de la muestra. Estos datos fueron usados para derivar un grupo de patrones de entrenamiento para la red. A las variables de entrada le fueron asignados valores numéricos entre 0 y 1 y la red fue entrenada para cada estudiante. El algoritmo de entrenamiento modificó iterativamente los valores numéricos de los "pesos" de conexión entre las neuronas para disminuir el error de entrenamiento de la red. Una vez entrenada la red, se fijaron los "pesos" definitivos. Cuando un grupo de datos se presentó a la red entrenada, se generaba un valor de salida representativo de la clasificación (éxito o fracaso) de la predicción. Este valor de salida se basa en el conocimiento que la red adquirió en la etapa de entrenamiento.

La red fue entrenada con un algoritmo de retropropagación para entrenamiento supervisado (Rumelhart y cil., 1986). El entrenamiento supervisado es una técnica en la que un grupo representativo de patrones entrada-salida es presentado a la red. A través de un algoritmo iterativo, los "pesos" internos de la red son ajustados para disminuir la diferencia entre la predicción proporcionalada por la red y el resultado verdadero de los casos de entrenamiento.

Una vez diseñado el modelo predictivo, se procedió a someterlo a una prueba de eficiencia. Se aplicó el modelo a la otra mitad de estudiantes de la muestra, es decir, datos que no habían intervenido en la construcción de la red. Este resultado es susceptible de ser comparado con el desempeño realmente obtenido por cada estudiante. Este caso se realizó por medio de un análisis ROC (Receiver Operating Characteristic), siguiendo el método descrito por Metznerll (1994).

La implementación de los algoritmos, así como el entrenamiento y examen de la red fueron realizados en lenguaje MATL.AB/SUN en un computador compatible con IBM con un procesador 486, 66 MHz.

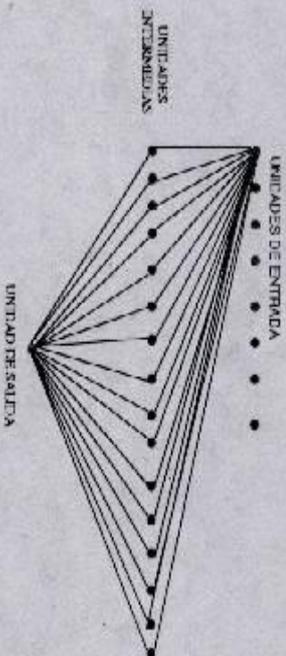


Figura 1.- Red neuronal artificial (8X 16X1) con 8 unidades de entrada, una capa de 16 unidades intermedias y una unidad de salida.

RESULTADOS

Después de una serie de ensayos, los parámetros de entrenamiento fueron optimizados así: número de unidades en la capa intermedia = 16; coeficiente de aprendizaje para las conexiones de la capa de entrada con la capa intermedia = 0.5; coeficiente de aprendizaje para las conexiones entre la capa intermedia y la neurona de salida=0.3; coeficiente de momento=1.0; e intervalo de entrenamiento=200 iteraciones. Todos los patrones fueron

presentados a la red en cada iteración. El tiempo de entrenamiento usualmente fue de 10 minutos para 200 iteraciones en todos los casos. Una vez entrenada la red, la evaluación de todos los casos requirió menos de 1 segundo.

El modelo predictivo se comportó satisfactoriamente al ser aplicado al pronóstico individual de 110 estudiantes (Tabla 1). La red predijo correctamente el rendimiento académico en 104 de los 110 casos presentados (94.55% de exactitud), en tanto que, según el análisis ROC, la sensibilidad y la especificidad del procedimiento fueron, respectivamente, de 92% y 96%. Dos casos clasificados como $RG=1$ por la red eran en realidad $RG=0$, y cuatro casos clasificados por la red como $RG=0$, eran en realidad $RG=1$. La red examina los 110 casos con un error de 0.0377 por estudiante.

TABLE 1. Comparación de la predicción del rendimiento académico con los resultados reales:

	RG = 1	RG = 0
TOTAL		
MODELO 110	74	36
REALES 110	76	34

DISCUSIÓN

El presente trabajo representaría una investigación inicial de la predicción del rendimiento académico con el uso de novedosas técnicas analíticas. En ese sentido, puede considerarse como una muestra de las posibilidades que ofrecen las técnicas de análisis de datos en las investigaciones sobre el rendimiento académico. El resultado obtenido demuestra ante todo la factibilidad de predecir, de una manera razonablemente confiable, el rendimiento académico a una escala individual, y al mismo tiempo, confirma los resultados de otras investigaciones (Borham y col., 1988; Lipson y col., 1988; Nnodim, 1994) que han demostrado que con el conocimiento de las características de los estudiantes en el momento de su ingreso en la universidad

se pueden construir herramientas con las que es posible predecir el rendimiento académico. Este pronóstico puede conseguirse al inicio del año académico, con información de fácil obtención. Si este pronóstico inicial, que separa a los estudiantes en dos grandes grupos, se complementa con las evaluaciones sistemáticas, se puede construir un perfil individual que describe la trayectoria del estudiante desde su ingreso hasta el final del curso. Según esta trayectoria y comprender las acciones, cuando sean requeridas, para corregirla, es un procedimiento que podría evitar sorpresas al final del curso.

Aunque este resultado no proporcionara una prueba definitiva de las ventajas de las redes neuronales artificiales sobre los sistemas tradicionales de predicción, un examen de las similitudes y diferencias entre ambos métodos puede ser útil para elegir el real potencial de aplicación de las redes neuronales. Superficialmente, una red neuronal artificial parece ser similar a otros algoritmos derivados estadísticamente en los que con un conjunto de datos de entrada se obtiene un resultado de salida. En la red neuronal, sin embargo, la relación entre las variables de entrada y las de salida no está definida ni cuantificada explícitamente. Por otra parte, el modelo de las redes neuronales difiere de los análisis estándares de predicción porque mantiene y procesa toda la información disponible en la base de datos; es decir, no descartaría variables que pueden ser críticas para la predicción. En nuestro estudio, conservando todas las variables, la red demostró su capacidad para manejar la heterogeneidad encontrada en la muestra de estudiantes. En este sentido, es conveniente tener presente que los métodos que asumen igual contribución de todas las variables exploratorias son más sencillos de desarrollar y proporcionan resultados comparables, en validez predictiva, a los obtenidos con modelos complejos de regresión (Wainer y Thissen, 1986). Finalmente, y aunque sólo puede ser postulado, uno de los aspectos más importantes es la capacidad de la red neuronal artificial de descubrir relaciones multidimensionales entre las variables lo que le permite identificar información y capitalizarla para mejorar su exactitud predictiva (Wasserman, 1990).

En el aspecto metodológico, puede concluirse que el modelo de la red neuronal artificial cubre las expectativas en el depositadas. La exactitud puede aún ser mejorada agregando otras variables que rutinariamente no son incluidas en los expedientes de ingreso. Es de esperar que, en la medida en que la base de datos se vuelva más grande, la red ganará más experiencia, hecho que le permitirá predecir con mayor exactitud. Sin embargo, la interpretación de los resultados no debe ser limitada, al menos en dos aspectos. En primer lugar, hasta que la red no sea examinada con un número mayor de estancias, el estudio debe ser visto como una primera etapa en la validación del uso de redes neuronales artificiales para la predicción del rendimiento académico y, en segundo lugar, la validación de la red requiere demostrar que es transportable, con el mismo nivel de exactitud, a diferentes grupos de datos.

BIBLIOGRAFÍA

- Bacalao, J. (1985). Divergos enfoques estadísticos del pronóstico del rendimiento académico. *Revista Cubana de Educación Superior*, 2: 27-34.
- Baxt, N.G. y Skora, J. (1996). Prospective validation of artificial neural network trained to identify acute myocardial infarction. *Lancet*, 347: 12-15.
- Borham, N.C., Russell, C. y Wastell, D.G. (1988). Can past academic criteria predict student at risk of future failure? *Medical Education*, 22: 94-98.
- Canavelli, P. (1995). Analysing the past and managing the future using neural networks. *Futures*, 27: 325-338.
- Khan, Z. H., Mohapatra, S. K., Khoshaf, P. K. y Raga Kumar, S. N. (1995). Artificial neural network and medicine. *Indian Journal of Physiology and Pharmacology*, 42: 321-342.
- Lipson, A., Hurham, G. y Hamilton, D. (1985). School results as predictors of medical school achievement. *Medical Education*, 22: 381-388.
- Meisner, M. L. (1990). Evaluation of neural network performance by receiver operating characteristic (ROC) analysis: examples from the biotechnology domain. *Comput Methods Programs Biomedical*, 32: 73-80.
- Miodini, J. O. (1994). Predicting performance in anatomy using antecedent academic criteria. *Medical Education*, 28: 200-207.
- Reichgelt, H. (1996). Neural networks in the study of the brain. *Molecular and Chemistry Neurophysiology*, 28: 231-235.
- Rumelhart, D. E., Hinton, G. E. y Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, 323: 533-536.
- Schmajda, N. A. y Thiemé, A. D. (1992). Purposive behavior and cognitive mapping: a neural network model. *Biological Cybernetics*, 67: 165-174.
- Stevens, R. H., Lopa, A. C. y Wang, F. (1990). Artificial neural networks can distinguish novice and expert strategies during complex problem solving. *Journal of American Medicine Information Association*, 3: 131-138.
- Warner, H. y Thiessen, D. (1986). Three steps towards robust regression. *Psychometrika*, 41: 9-34.
- Wasserman, P. D. (1990). *Neural computing: theory and practice*. New York: Van Nostrand Reinhold.
- Yee, D., Prior, M. G. y Florence, L. Z. (1993). Development of predictive models of laboratory animal growth using artificial neural networks. *Comput Applied Bioscience* 9: 517-522.

