

Aplicación de Redes Neuronales Artificiales en la Orientación Vocacional

Romero Aquino, Nelson Marcelo; Martínez Jara, Eustaquio Alcides
Universidad Nacional del Este, Facultad Politécnica

Resumen

Implementando técnicas de inteligencia artificial se desarrolla un sistema con la capacidad de generar recomendaciones vocacionales que puedan proporcionar soluciones al problema de selección de profesión por el que atraviesan los jóvenes al finalizar sus estudios secundarios. En un principio, un grupo de individuos compuesto por alumnos del Colegio Sembrador y del Centro Regional de Educación son sometidos a pruebas de orientación vocacional que se ponen a disposición en una página web. Posteriormente se determinan las áreas vocacionales recomendadas para cada sujeto con asistencia de un profesional. Finalmente, tanto los resultados de los tests como las recomendaciones profesionales son utilizados como datos de ejemplo para someter al sistema a un proceso de aprendizaje en el que adquiere, mediante algoritmos de inteligencia computacional, la capacidad para ofrecer recomendaciones de forma autónoma tomando únicamente como referencia, en operaciones posteriores, los resultados de las pruebas vocacionales.

Descriptor: Redes neuronales artificiales, Inteligencia artificial, Selección de personal, Orientación vocacional.

Abstract

In order to solve the profession selection problem, a problem that affects most of the people at the time they finish their high school, artificial intelligence techniques are implemented in the development of a system capable to generate vocational recommendations. At first, a group of students from the Colegio Sembrador and the Centro Regional de Educación complete vocational tests, which are available online

in a web page. Next, with the assistance of a professional, one or more vocational areas are recommended for each individual. Finally, the results of the tests and the professional recommendations are used as samples data in a procedure where the system learns and acquires, by using computational intelligence algorithms, the capability to provide recommendations automatically in future operations using as its input data nothing else but the results of the vocational tests.

Keywords: Artificial neural networks, Artificial intelligence, Personnel selection, Vocational Orientation.

1. Introducción

Para afrontar la necesidad de seleccionar una carrera profesional durante los últimos años de la etapa escolar, las pruebas de orientación vocacional se constituyen en herramientas útiles y muy utilizadas. Sus resultados ofrecen datos que relacionan a una persona con distintas áreas y le indican su nivel de pertenencia a las mismas. Si bien la mayoría de las pruebas vocacionales proporcionan resultados fácilmente interpretables por los individuos (Anónimo, 2014), existen otras cuya interpretación es posible únicamente con la asistencia de un profesional o tras un exhaustivo análisis y estudio de los distintos aspectos que enumera el test y la relación de dichos aspectos con las diversas profesiones que el individuo tiene como opción (Thomas, 2014). Una gran problemática debido a que en ocasiones no es posible contar con la asistencia del especialista.

Las Redes Neuronales Artificiales son un paradigma de aprendizaje y procesamiento automático, basado en el funcionamiento del sistema nervioso de los seres vivos, y cuyo objetivo es el procesamiento de información para la resolución de problemas (Wikipedia, 2014).

La principal característica de las RNA es su capacidad para construir un sistema capaz de resolver problemas de clasificación y predicción de patrones tras un proceso de aprendizaje en el cual recibe datos de ejemplo basados en soluciones para el problema que se pretende modelar, y que se han obtenido previamente. No son necesarios algoritmos o modelos definidos previamente para que una Red

Neuronal encuentre la solución a un problema, sólo son necesarios los datos de ejemplo.

Teniendo en cuenta dicha característica, el presente trabajo plantea la implementación de una Red Neuronal Artificial capaz de recomendar áreas vocacionales en base a las informaciones obtenidas en pruebas de orientación vocacional, teniendo como principal objetivo solucionar la problemática consistente en que el joven que finaliza sus estudios secundarios no cuenta con la posibilidad de interpretar sus resultados obtenidos en pruebas vocacionales, o tras combinar los resultados de dos o más pruebas.

Se propone la implementación de las Redes Neuronales debido a que el problema de recomendación de áreas vocacionales no puede ser programado tradicionalmente (al tratarse las recomendaciones de un conjunto de opiniones subjetivas que naturalmente no comprenden un modelo identificable) sino que sólo se cuenta con datos de ejemplos (las recomendaciones vocacionales para un cierto conjunto de resultados de un individuo en un test vocacional) que el sistema deberá ser capaz de comprender y emular (Wikipedia, 2014).

1.1. Motivación

- Las Redes Neuronales Artificiales tienen una amplia gama de aplicaciones.
- Se pueden utilizar en problemas en donde se cuenta únicamente con valores de entrada/salida.
- La IA y la Psicología han sido relacionadas en contadas ocasiones, es un área en expansión.
- La posibilidad de realizar una comparación entre el método de orientación vocacional a través de RNAs y los métodos psicológicos tradicionales.

1.2. Objetivos

1.2.1. General

Implementar una red neuronal artificial capaz de determinar la profesión de un individuo en base a sus aptitudes y preferencias.

1.2.2. Específicos

- Reunir información acerca de aptitudes y preferencias de individuos a través de pruebas psicológicas.
- Modelar vectores de entrada y salida (datos de entrenamiento) a partir de los datos recabados en las pruebas psicológicas.
- Implementar la red neuronal artificial y establecer la configuración más eficiente de la misma en base a la experimentación.
- Analizar los resultados obtenidos y el porcentaje de acierto de la red.

1.3. Literatura Relevante

La orientación vocacional está asociada al área de la selección de personal y existen trabajos investigativos que han buscado resolver problemas de dicha índole mediante la implementación de redes neuronales artificiales. A continuación se exponen dos de ellos.

1.3.1. Selección de personal mediante redes neuronales artificiales

El proyecto de "Selección de personal mediante redes neuronales artificiales" fue realizado en el 2006, en la Escuela de Ingeniería Eléctrica y Electrónica de la Universidad del Valle, el trabajo consistió en seleccionar candidatos que fueran aptos para ingresar a la Escuela Naval Almirante Padilla de Cartagena, Colombia. El sistema estaba basado en la prueba psicológica 16-PF, de donde se extraen los factores de Primer Orden y la distorsión motivacional, para ser utilizados como entradas a un algoritmo de reconocimiento de patrones que actúa como predictora de comportamiento. El experimento utilizó dos tipos de redes neuronales artificiales, en donde el Perceptrón Multicapa arrojó una efectividad del 60% y la Red de Base Radial obtuvo 70% de eficiencia (Acevedo, Caicedo & Loaiza, 2007).

1.3.2. Aplicación de una Red Neuronal para clasificación de perfiles psicológicos con el test 16PF del Dr. R.B. Cattell en una población estudiantil.

El objetivo de este trabajo, realizado por el Ing. Adolfo Gordillo en Colombia, era el de detectar posibles rasgos conflictivos de personalidad en estudiantes de tal

forma a aplicar el tratamiento pertinente de acuerdo a los resultados arrojados en el procesamiento. En el trabajo fue utilizada la prueba 16-PF y la red neuronal, un Perceptrón Multicapa, obtuvo una efectividad casi del 100% (Gordillo, 2014).

2. Marco Teórico

2.1. Inventario de Intereses de Hereford

Cuestionario de 90 preguntas que son respondidas numéricamente seleccionando un número del 1 al 5. Cada pregunta forma parte de una determinada área de interés, por ejemplo "tocar un instrumento musical" forma parte del área Musical, mientras que "Leer a los clásicos" es una pregunta referente al área de Literatura.

2.2. Test de Aptitudes Diferenciales (DAT)

Está constituida por 8 formularios que contienen distintos tipos de ejercicios, cuya cantidad varía en cada caso. A diferencia del inventario de intereses de Hereford, en el que no hay resultados incorrectos sino que sólo estimaciones de grados de interés, las respuestas a los ejercicios de los formularios del test DAT pueden estar correctas o incorrectas (Thomas, 2014).

2.3. HTML (HyperText Markup Language)

Es un lenguaje de marcado utilizado para la creación de páginas web. Permite describir documentos de texto presentándolos de forma estructurada, con enlaces o hyperlinks que conducen a otros documentos o fuentes de información, como bases de datos, que pueden encontrarse almacenados en una máquina local o en máquinas remotas de la red. Además de textos, también es posible incluir medios multimedia como sonido, imágenes, gráficos estadísticos, etc., de acuerdo a lo que se requiera (The PHP Group, 2014).

2.4. PHP (Hypertext Preprocessor)

Es un lenguaje de propósito general y de código abierto especialmente adecuado para el desarrollo web y que puede ser incrustado en páginas HTML. El código PHP se ejecuta del lado del servidor, en lugar del navegador, con un módulo de procesador de PHP que genera la página web. Su sintaxis es similar a C, Java y Perl (Wikipedia, 2014).

2.5. Redes Neuronales Artificiales

Las Redes Neuronales Artificiales (RNA) son un sistema cuya arquitectura se basa en una interconexión de neuronas artificiales, o pequeñas unidades de procesamiento, dispuestas de forma paralela en una red (Wikipedia, 2014).

Un proceso de aprendizaje ocurre durante el funcionamiento y procesamiento de la red, por medio de un procedimiento en el que las conexiones entre las neuronas artificiales son actualizadas, son en éstas conexiones en donde reside realmente la inteligencia de la red (Wikipedia, 2014).

Tras el proceso de aprendizaje, el margen de error o la tasa de acierto del sistema es calculado por la red, previamente separando un porcentaje del total de muestras, compuestas por valores de entrada para la red y valores de salida deseada de la misma, para utilizarlos en el proceso de prueba, el cual consiste en llevar a cabo una comparación entre dicha salida deseada con la salida real provista por la red (Russell & Norvig, 2004).

El tipo de red neuronal utilizado para el desarrollo del trabajo ha sido el Perceptrón Multicapa (MLP), un tipo de red neuronal que incorpora más de una capa de unidades de procesamiento (Palacios, 2014), considerado como un aproximador universal. Actualmente es una de las arquitecturas de red neuronal más utilizadas.

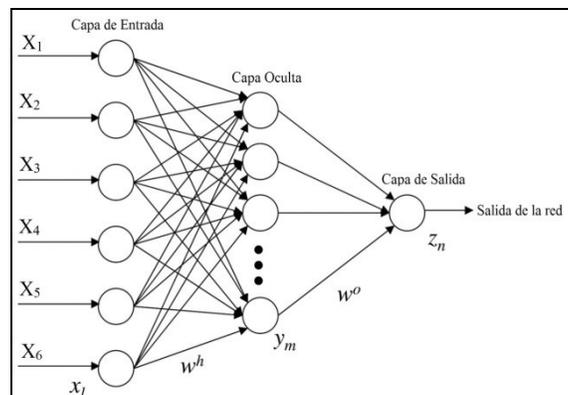


Figura 1. Modelo de MLP

3. Materiales y Métodos

3.1. Materiales

3.1.1. Hardware

- Computador
 - Procesador de 4 núcleos corriendo a 3.5 GHz (64 bits)
 - Memoria RAM de 8 Gb DDR3
 - Disco Duro de 1 Tb

3.1.2. Software

- MatLab versión 7.10.0 para la implementación de la red neuronal
- Adobe Dreamweaver CC para el desarrollo de la página web.
Especificaciones del servidor en el que se ha alojado la página web:
 - Servidor Apache versión 2.2.25, con hosting compartido, administrado vía cPanel 11.4.
 - PHP versión 2.2.25.
 - Base de datos MySQL versión 5.1.70.
 - Mecanismo de almacenamiento de datos InnoDB transaccional.
 - Sistema operativo Red Hat Linux versión 5.10.
 - Kernell del S.O. versión 2.6.18.
 - CPU multiprocesador de 2 procesadores con 4 núcleos cada uno.

3.2. Metodología

Para el desarrollo del trabajo se utilizaron dos pruebas vocacionales: el Inventario de Intereses de Hereford y el Test de Aptitudes Diferenciales (DAT). Ambos test se han puesto a disposición en una página web programada en HTML y PHP.

Un total de 78 alumnos del nivel medio del Colegio Sembrador y del Centro Regional de Educación Dr. José Gaspar Rodríguez de Francia completaron ambas pruebas vocacionales accediendo a la página web.

Una vez reunidas todas las muestras, consistentes en los resultados de los alumnos en las dos pruebas vocacionales, se llevaron a cabo las recomendaciones vocacionales y, posteriormente, la red neuronal artificial fue implementada.

La estructura de datos establecida para el tratamiento de los mismos, tanto para efectuar la recomendación como para implementar la red neuronal ha sido la siguiente:

- **Datos de entrada:** resultados de las pruebas vocacionales, los cuales comprenden 17 ítems (9 resultados del test Hereford y 8 resultados del test DAT) normalizados que pueden tomar valores numéricos comprendidos entre 0 y 1.
- **Datos de salida:** recomendaciones vocacionales compuesto por cuatro ítems binarios, que pueden tomar valores 0 y 1. Cada ítem representa un área vocacional (Física/Matemática, Químico/Biológica, Ciencias Sociales, Humanidades/Artes). El valor de salida del ítem determina si el área vocacional al que éste hace referencia es recomendada (valor 1) o no (valor 0).

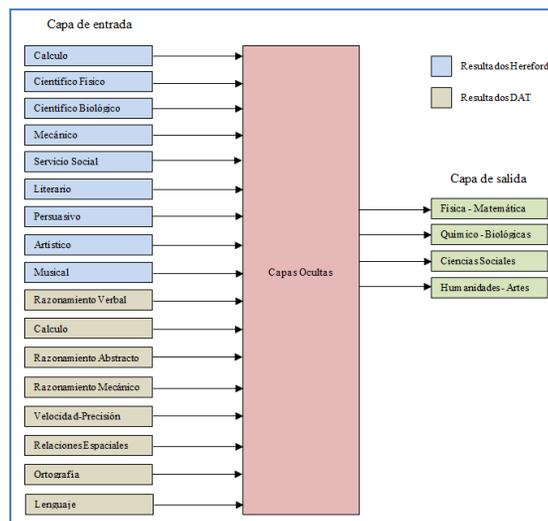


Figura 2. Diseño A de la red implementada

Los datos de entrada se han diseñado de dos formas, y la RNA se ha implementado utilizando ambas estructuras, la primera se puede observar en la Figura 2, que se constituye en el diseño señalado anteriormente.

El segundo diseño se puede ver en la Figura 3, se trata de una estructura en la que se obvian los resultados del test Hereford y sólo se utilizan los del test DAT como valores de entrada.

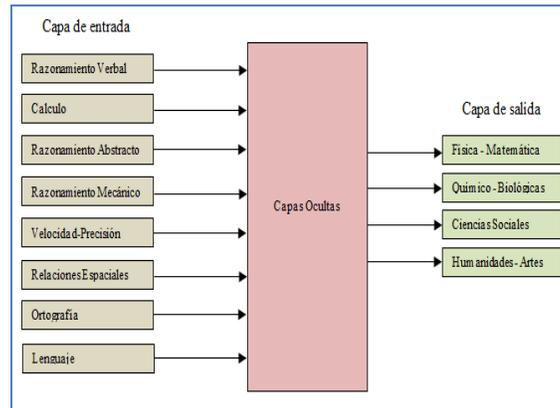


Figura 3. Diseño B de la red implementada

Una vez que se han hecho todas las recomendaciones correspondientes para cada sujeto, dividieron las muestras en dos grupos: 60% para el entrenamiento de la red y 40% para la prueba.

La red neuronal artificial se ha implementado utilizando la herramienta MatLab para las dos estructuras de datos de entrada vistas con anterioridad, y se ha ejecutado una etapa de experimentación en la que se ha buscado una configuración de red que otorgara los resultados o recomendaciones que fueran lo más aproximados posibles a las recomendaciones hechas con asistencia del especialista, lo cual se puede comprobar contrastando la salida de la red tras procesar los datos de prueba (40% de las muestras) con la salida esperada para esos datos de prueba.

Durante la experimentación se han ido alterando los parámetros estructurales (cantidad de neuronas y de capas ocultas) y los de entrenamiento (objetivo, cantidad de épocas, etc.) de las redes neuronales implementadas, en la búsqueda de resultados esperados.

Para comprobar la eficiencia de las distintas configuraciones de red en el proceso de prueba, se han utilizado gráficos de curvas ROC. Las mismas representan la sensibilidad (razón de verdaderos positivos o VPR) frente a $(1 - \text{especificidad})$,

siendo la *especificidad* la razón de falsos positivos o FPR. La VPR establece el grado de efectividad que posee el clasificador para detectar o clasificar los casos positivos correctamente, de entre todos los casos positivos disponibles para la prueba. Mientras que la FPR estima la cantidad de valores positivos incorrectos que produce el clasificador, cuando estos deberían ser negativos, comparando los valores positivos con la cantidad total de negativos disponibles para la prueba.

Se considera que la red tiene una gran eficiencia y se constituye en un buen clasificador cuando el punto medio de todas las curvas ROC se encuentran más alejados del eje x (mayor tasa de verdaderos positivos) y más cerca del eje y (menor tasa de falsos positivos). Por otra parte, cuando una o más curvas se encuentran muy cerca de la recta diagonal en donde Tasa de Falsos Positivos = Tasa de Verdaderos Positivos, la red sólo es capaz de realizar una clasificación aleatoria; mientras que cuando una o más curvas se encuentran por debajo de dicha recta, la red neuronal es un mal clasificador.

4. Resultados

Las distintas configuraciones de redes neuronales en donde se han utilizado dos tests vocacionales han sido descartadas tras la experimentación, por no lograr una eficiencia aceptable tras ser entrenadas, comprobable en el proceso de prueba utilizando el gráfico de curvas ROC, como se puede observar en la Figura 4, correspondiente a una configuración de red en la que se utilizan dos tests vocacionales.

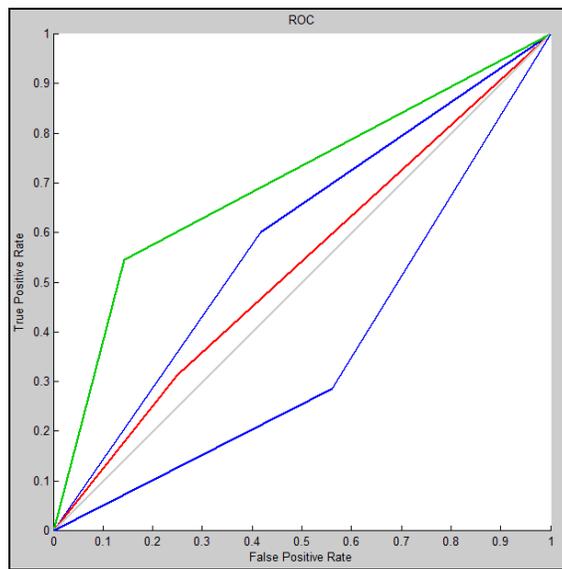


Figura 4. Gráfico ROC de red entrenada utilizando dos pruebas vocacionales.

La red neuronal con los mejores resultados es una de las que tienen como valores de entrada los resultados de un solo test: el Test de Aptitudes Diferenciales (DAT). Su configuración es la siguiente:

Neuronas en la capa de entrada: 8

Neuronas en la primera capa oculta: 10

Función de activación en la primera capa oculta: sigmoide (logsig)

Neuronas en la segunda capa oculta: 10

Función de activación en la segunda capa oculta: sigmoide (logsig)

Neuronas en la capa de salida: 4

Función de activación en la capa de salida: sigmoide (logsig)

Sus parámetros de entrenamiento son:

Epochs: 1000 iteraciones

Min_grad: $1,00 \times 10^{-10}$

Goal: $1,00 \times 10^{-10}$

Mu_max: $1,00 \times 10^3$

Y su curva ROC resultante de los datos de prueba se puede observar en la Figura 5.

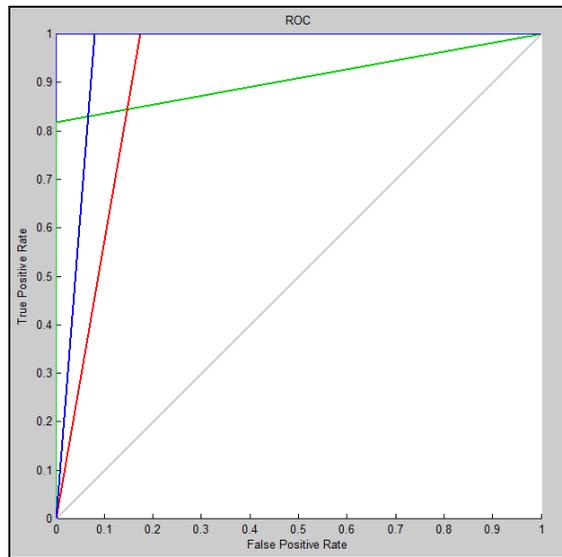


Figura 5. Gráfico ROC de la red entrenada con la mejor configuración.

El algoritmo de aprendizaje para todas las redes neuronales implementadas fue el de Levenberg-Marquardt.

5. Conclusión

Teniendo en cuenta los resultados obtenidos, es posible modelar y simular la subjetividad a través de la implementación de redes neuronales artificiales, entendiéndose a la subjetividad como una opinión o recomendación que tiene como base una apreciación personal de un especialista o que es realizada con asistencia del mismo.

Se ha podido recrear un sistema capaz de aprender a efectuar recomendaciones con un alto grado de acierto con respecto a las que son realizadas con asistencia profesional.

Sin embargo, el modelado de los datos que son tratados por la red debe hacerse de tal modo que los valores de entrada estén estrechamente relacionados entre sí y que exista un patrón de salida determinado que pueda ser reconocido por la red durante su proceso de entrenamiento, como ha ocurrido cuando la red ha sido entrenada utilizando solamente como valores de entrada los resultados del test DAT.

En caso contrario, ocurrirá lo que ha sucedido cuando se implementaron los resultados tanto del test DAT como del test Hereford como valores de entrada de la red, en donde no existía cohesión entre los mismos, al ser éstas estimaciones muy dispares e incluso contradictorias (la primera ofrece estimaciones acerca de las aptitudes mientras que la otra lo hace acerca de los intereses), y que su combinación no desemboque en la existencia de un patrón, es decir, que los datos de entrada y sus resultados sean aleatorios, lo cual resulta en la imposibilidad de que la red aprenda a pesar de las distintas configuraciones que ésta pueda tomar.

Referencias Bibliográficas

Acevedo, G., Caicedo, E. & Loaiza, H. (2007). Selección de personal mediante redes neuronales artificiales. *Revista de Matemática: Teoría y Aplicaciones*, 14(1), 7-20.

Anónimo. (2014). Test de Orientación Vocacional (CHASIDE). Recuperado de <http://www.tecnicas-de-estudio.org/general/test-orientacion.htm>

Gordillo, A. (2014). Aplicación de una Red Neuronal para clasificación de perfiles psicológicos con el test 16PF del Dr. R. B. Cattell en una población estudiantil. Recuperado de <http://espejos.unesco.org.uy/simplac2002/Ponencias/Inforedu/IE004%20Adolfo%20A.%20Gordillo.doc>

Palacios, F. (2014). Perceptrón Multicapa. Recuperado de http://www.ibiblio.org/pub/linux/docs/LuCaS/Presentaciones/200304curso-glisa/redes_neuronales/curso-glisa-redes_neuronales-html/x105.html

Russell, S., Norvig, P. (2004). *Inteligencia Artificial, Un enfoque moderno*. Madrid: Prentice Hall.

The PHP Group. (2014). Manual de PHP – Prefacio. Recuperado de <http://www.php.net/manual/es/preface.php>

Thomas, U. (2014). Test DAT – Test de Aptitudes Diferenciales. Recuperado de <http://www.elpsicoasesor.com/2011/03/test-dat-test-de-aptitudes.html>

Wikipedia. (2014). Redes Neuronales Artificiales. Recuperado de http://es.wikipedia.org/wiki/Red_neuronal_artificial

Wikipedia. (2014). Servidor HTTP Apache. Recuperado de http://es.wikipedia.org/wiki/Servidor_HTTP_Apache