



Modelización, predicción y clasificación del rendimiento académico de los alumnos mediante redes neuronales artificiales

E. T. Lau¹ · L. Sun^{1,2} · Q. Yang¹

© El autor(es) 2019

ABIERTO

Abstracto

Las evaluaciones estadísticas convencionales se limitan a proporcionar buenas predicciones de la calidad educativa universitaria. Este artículo presenta un enfoque con análisis estadístico convencional y modelado/predicción de redes neuronales del desempeño de los estudiantes. Las evaluaciones estadísticas convencionales se utilizan para identificar los factores que probablemente afecten el desempeño de los estudiantes. La red neuronal se modela con 11 variables de entrada, dos capas de neuronas ocultas y una capa de salida. El algoritmo de Levenberg-Marquardt se emplea como regla de entrenamiento de retropropagación. El rendimiento del modelo de red neuronal se evalúa a través del rendimiento de error, la regresión, el histograma de error, la matriz de confusión y el área bajo la curva de características operativas del receptor. En general, el modelo de red neuronal ha logrado una buena precisión de predicción del 84,8 %, junto con limitaciones.

Palabras clave Rendimiento académico · Análisis estadístico · Red neuronal artificial · Aprendizaje automático

1. Introducción

El rendimiento académico de los estudiantes universitarios es el referente más importante para comparar la calidad de los estudiantes universitarios. Sirve como criterio básico para que la universidad controle la calidad de la enseñanza y el aprendizaje y para que la universidad evalúe y seleccione a los estudiantes. Hoy en día, la mayoría de las universidades enfrentan desafíos difíciles para atraer a futuros estudiantes debido a mercados educativos cada vez más competitivos. Por lo tanto, el estudio de los logros académicos de los estudiantes es de gran importancia para promover el desarrollo de los estudiantes y la mejora de la calidad de la educación superior. Sin embargo, el rendimiento del estudiante se ve influido por muchos factores de manera complicada, y el entorno socioeconómico del estudiante y su rendimiento académico histórico pueden afectar potencialmente su rendimiento académico. Como era de esperar, la mayoría de los trabajos de investigación existentes se han limitado a analizar y predecir el desempeño de los estudiantes en la formulación de un problema relativamente simple utilizando técnicas estadísticas.

Para hacer frente a tal limitación, el aprendizaje automático se ha utilizado cada vez más en las aplicaciones de ciencia de datos para analizar relaciones complejas. Es capaz de aprender automáticamente sin ser programado explícitamente. Un modelo de red neuronal artificial (ANN), a pesar de que tiene una larga historia establecida en informática y ciencia de datos, está ganando una atención creciente y amplias aplicaciones. ANN amplía la capacidad de analizar una cantidad complicada de conjuntos de datos que no se pueden simplificar fácilmente mediante las técnicas estadísticas convencionales. También tiene la capacidad de detectar implícitamente relaciones no lineales entre variables dependientes e independientes [1]. ANN ha ganado una mayor atención y ha demostrado ser un gran éxito en la aplicación de reconocimiento de patrones, clasificaciones, pronósticos y predicciones en las áreas de atención médica, clima y tiempo, mercados bursátiles, etc.

Sin embargo, el uso de ANN es limitado en el área de investigación educativa. Esto puede deberse a la complejidad que surge de la red modelada, la dificultad de un sistema ANN modelado para proporcionar una explicación adecuada (la caja negra).

* ET Lau, engtseng.lau@brunel.ac.uk; L. Sun, sunlizhi@qfnu.edu.cn; P. Yang, qingping.yang@brunel.ac.uk | Kingston Lane, Uxbridge UB8 3PH, Reino Unido.

²Normal University, Qufu 273165, Shandong, República Popular de China.

¹Brunel University London, Qufu



naturaleza), la propensión al ajuste excesivo y el tiempo necesario para el entrenamiento de la red neuronal [1]. Para mitigar las deficiencias a través de la aplicación de ANN, en este documento se compromete un dominio de competencia tanto para el análisis estadístico convencional como para el ANN. Los datos educativos se estudiarán inicialmente con análisis estadístico convencional, y la confirmación de los resultados estadísticos se utilizará para realizar el entrenamiento, la validación y las pruebas de ANN con el fin de desarrollar el modelo ANN con ajustes de configuración adecuados para predecir y clasificar con precisión el rendimiento de los estudiantes. .

En general, este documento presenta un modelo exploratorio y análisis del desempeño de los estudiantes a través de los datos recopilados de una universidad china. El modelo ANN sirve como una herramienta de calidad educativa dominante que evalúa el desempeño de los estudiantes en todas las universidades, aborda esas disparidades y, por lo tanto, mejora continuamente la calidad de la educación.

La organización del trabajo es la siguiente: Secc. 2 presenta las revisiones de ANN y su resumen, Secc. 3 presenta la metodología de pruebas estadísticas, modelado y verificación de ANN, Secc. 4 presenta los resultados de las evaluaciones estadísticas, los ajustes de configuración de ANN y el rendimiento, Secc. 5 concluye los hallazgos.

2 Red neuronal artificial

La red neuronal artificial (ANN) es una herramienta de modelado poderosa y compleja para modelar funciones no lineales que a menudo describen los sistemas del mundo real [2–4]. ANN se forma a través de una colección de neuronas artificiales que se asemejan a la geometría de conexión de las neuronas en el cerebro humano para ejecutar una tarea con un rendimiento mejorado a través del "aprendizaje, la capacitación y la mejora continua" [2, 5, 6].

Las RNA se forman con estructuras neuronales de tres capas, a saber, las capas de entrada, oculta (media) y de salida. La capa de entrada recopila datos de información numérica con conjuntos de características y valores de activación. Los valores de entrada se propagan a través de las neuronas interconectadas a la capa oculta. En la capa oculta, las neuronas de entrada se suman para calcular la suma ponderada de las neuronas de entrada; y las neuronas sumadas se combinan para producir resultados en la capa de salida usando una función de activación (o transferencia) [5, 7]. Tanto las neuronas como la conexión contienen pesos ajustables durante el proceso de aprendizaje. Las neuronas sumadas se transformarán matemáticamente en la capa de salida si se supera el umbral de la función de activación.

La cantidad de veces que se utilizan las funciones de entrenamiento para actualizar los pesos de conexión en el proceso de alimentar los valores de entrada y terminar con los valores de salida en ANN se denomina Época [7]. Aquí es donde las entradas

de neuronas artificiales se multiplican por pesos, y el resultado de estas sumas se envía a la capa de salida a través de una función de activación [6]. Las funciones de activación más utilizadas incluyen funciones tangentes lineales, sigmoideas e hiperbólicas. El entrenamiento termina cuando se alcanza el valor de época máximo y/o las comprobaciones de validación. Los datos entrenados resultantes se introducen en los datos de prueba para examinar el rendimiento de la ANN.

La regla de aprendizaje más común de las ANN es la retropropagación (BP), que es un enfoque de aprendizaje supervisado y se puede utilizar para entrenar las redes neuronales profundas. BP ajusta los pesos de las neuronas a través de los errores calculados y permite que la red aprenda del proceso de entrenamiento. La resolución típica de problemas de las ANN incluye tres arquetipos de aprendizaje, es decir, aprendizaje supervisado, aprendizaje no supervisado y aprendizaje por refuerzo [6].

Sorprendentemente, el enfoque ANN ha recibido una gran atención con fines de investigación educativa. [2] aplicó ANN para modelar y realizar entrenamiento de datos basado en el comportamiento de selección de cursos de los estudiantes, y además identificó la mejor estrategia y configuración para satisfacer las demandas de los estudiantes para cada curso para una programación de cursos óptima dentro de una universidad. ANN se utilizó para clasificar a los estudiantes dentro de la facultad musical para predecir las percepciones de los estudiantes en educación musical [4]. ANN se combinó con la optimización de enjambres de partículas para evaluar el rendimiento de los profesores y también para mejorar la precisión del reconocimiento en el sistema de evaluación de profesores de la universidad [7]. [8] propuso un procedimiento de inteligencia artificial (IA) basado en un modelo de red neuronal autoorganizada que caracterizó automáticamente los perfiles bibliométricos de los investigadores académicos e identificó aún más las instituciones que tenían un patrón similar de rendimiento académico. Además, [5] aplicó diferentes modelos de predicción además de la técnica ANN, como análisis discriminante, bosques aleatorios y árboles de decisión, para predecir y clasificar el rendimiento académico de los estudiantes en tres universidades diferentes.

ANN también se usó para modelar y simular la diversidad de estilos de aprendizaje entre los estudiantes a través de dos paradigmas de aprendizaje, a saber, el aprendizaje supervisado (aprendizaje con el maestro) y el aprendizaje no supervisado (aprendizaje sin maestro) y a través del autoaprendizaje de los estudiantes [9]. De manera similar, [10] propuso una herramienta basada en inteligencia artificial que tenía en cuenta ANN como uno de los métodos de inteligencia artificial para minimizar la desorientación del comportamiento de aprendizaje y los problemas cognitivos sobrecargados entre los estudiantes. En la herramienta, la ANN se desempeñó mejor en comparación con otros modelos de aprendizaje aplicados. Además, se aplicó una ANN basada en BP para evaluar la calidad del sistema de enseñanza y el rendimiento de la ANN pudo cumplir con los requisitos de viabilidad y precisión del sistema [3]. Aunque los modelos clasificaron con precisión a los estudiantes a

el proceso de predicción con diferentes niveles de calificaciones académicas, se obtuvo como resultado bajas precisiones de predicción. El artículo de [11] aplicó con éxito ANN para predecir el estado de ánimo del estudiante durante la prueba de autoevaluación en línea, con una precisión de predicción de más del 80%. [12] adoptó un enfoque similar, pero se centró en predecir el éxito académico en cursos de matemáticas utilizando ANN basado en BP en tres universidades diferentes, con una precisión de predicción de 93,02%.

En general, ANN tiene la capacidad de realizar el ajuste neuronal y la predicción, y la capacidad de clasificar teóricamente cualquier dato con precisión arbitraria. Sin embargo, solo existen estudios y herramientas limitados para predecir el rendimiento académico de los estudiantes, especialmente en el contexto de la educación inteligente. Por lo tanto, este artículo ha propuesto el uso de ANN como una herramienta de modelado de aplicaciones para predecir el rendimiento académico entre los estudiantes. A diferencia de los conceptos de modelado descritos en la literatura, el documento se centra en los antecedentes socioeconómicos de los estudiantes y los resultados de sus exámenes de ingreso que probablemente afectarían su desempeño académico general en una universidad. Esto es importante no solo para examinar los factores significativos que afectan el desempeño de los estudiantes, sino también para clasificar correctamente el desempeño de los estudiantes en función del patrón predicho obtenido de ANN. Dado que el rendimiento de la predicción puede verse muy reducido por las enormes discrepancias de las muestras de diferentes universidades [5], este documento solo se centrará en muestras de datos de una sola universidad para garantizar la precisión de la predicción del rendimiento académico. El modelo ANN modelado serviría como marco y herramienta para predecir el desempeño académico de los futuros estudiantes y para abordar aún más aquellos problemas que dificultan el éxito del aprendizaje de los estudiantes y, por lo tanto, mejorar continuamente la calidad educativa.

3. Metodología

Esta sección presenta la metodología de adquisición de datos educativos, pruebas de hipótesis estadísticas para el modelado de redes neuronales y la metodología de evaluación del rendimiento de ANN modeladas.

3.1 Recopilación de datos educativos

Los datos de la muestra se recopilaron sobre un total de 1000 estudiantes, 275 mujeres y 810 hombres, seleccionados entre los estudiantes de pregrado de la Universidad Q. Los resultados del examen comprendieron tres programas de pregrado dentro de tres departamentos de la universidad para la admisión de estudiantes desde el año 2011 hasta el 2011. 2013, con cuatro años en total de cursos por cada año de admisión. Además, los datos muestrales del entorno socioeconómico de los estudiantes

y también se recopilaron los resultados del examen de ingreso a la universidad nacional. Los resultados del examen de ingreso incluyeron cinco materias básicas: chino, inglés, matemáticas, ciencias integrales y prueba de competencia. El rendimiento de los estudiantes se determinó utilizando el Promedio de calificaciones acumulado estandarizado (CGPA) durante los cuatro años completos de duración de sus estudios. Se calcula como la media ponderada de la nota obtenida:

$$CGPA = \frac{\text{¿Cuál es el punto?}}{\text{allí}}, \quad (1)$$

donde c_i denota las horas de crédito por curso i y g_i es el punto de calificación recibido por curso i .

3.2 Prueba de hipótesis estadística

Previo al modelado ANN, es necesario examinar la relación entre el entorno socioeconómico, familiar y educativo de los estudiantes y su rendimiento académico. Las evaluaciones estadísticas se realizan inicialmente en este estudio, incluidas las correlaciones, la prueba t de dos muestras y ANOVA (factor único). Los coeficientes de correlación de Pearson se calculan para medir las relaciones lineales de cinco materias principales del examen de ingreso y el CGPA resultante. Esas variables correlacionadas se utilizan además como neuronas de entrada para el modelado ANN. Para la prueba t y ANOVA, se elige un nivel de significación de $p = 0,05$ para probar la diferencia significativa, en particular la diferencia media de CGPA en dos o más grupos de muestras. La prueba de hipótesis en este caso se realiza para comprobar si los estudiantes masculinos y femeninos, las ubicaciones geográficas, los roles de los padres y los tipos de estudiantes tienen una influencia significativa en el CGPA resultante.

3.3 Modelado de redes neuronales

ANN en este caso se usa para la predicción neuronal del CGPA de los estudiantes y la clasificación de datos a través de observaciones de entrada. Ambos esquemas se realizan en base a aprendizaje automático supervisado. La ANN en este documento se modela siguiendo los trabajos de investigación anteriores [5, 7, 13].

No es el alcance de este documento comparar la efectividad del rendimiento de ANN con otras técnicas de aprendizaje automático.

Por lo general, el modelo ANN se puede expresar simplemente como un función matemática:

$$= f(\vec{y}, \vec{y}), \quad (2)$$

donde \vec{y} y \vec{y} son los vectores de salida y entrada. \vec{y} es un vector de parámetros de peso que representan las conexiones dentro de la ANN.

La capa de entrada recopila datos con conjuntos de características y los valores de entrada se envían a la capa oculta. Los valores de salida

de la j -ésima neurona y_j del vector \vec{y} se calculan mediante la suma ponderada de los elementos de entrada x y w :

$$y_j = \sum_{i=1}^{N_i} w_{ij} x_i \quad (3)$$

La \vec{y} es la función de activación (función de transferencia), N_i es el número total de i -ésimas líneas de conexión a la j -ésima neurona y x_i es el valor de salida de la capa anterior de la i -ésima neurona. La función de activación (\vec{y}) de la tangente hiperbólica se usa para transferir el valor de la suma ponderada de las entradas a la capa de salida. Por lo tanto, el nodo activado resultante para la siguiente capa de entrada es:

$$x_j = \vec{y}(y_j) \quad (4)$$

Para reducir la dimensionalidad del espacio del predictor y evitar la posibilidad de sobreajuste, se emplea el análisis de componentes principales (PCA) [14]. PCA es una técnica de reducción de datos que transforma los predictores en forma temprana, elimina cualquier dimensión redundante y genera nuevos conjuntos de variables llamadas componentes principales [14].

El enfoque de aprendizaje supervisado basado en BP se aplica donde los parámetros de entrada y salida se suministran al modelo ANN. BP se usa como la regla de aprendizaje para el modelo ANN que ajusta los pesos de las neuronas w_{ij} a través de los errores calculados que además producen los resultados deseados. La función de error (E) de la ANN calculada basada en BP se calcula como la suma de la diferencia de la diferencia cuadrática entre los valores objetivo y los resultados deseados:

$$E = \frac{1}{2} \sum_j (\vec{y}_j - t_j)^2 \quad (5)$$

donde t_j es el valor objetivo para la neurona j en la capa de salida y N_j es el número total de neuronas de salida.

En ANN, los pesos se actualizan recursivamente. El algoritmo de optimización Levenberg-Marquardt basado en BP se aplica en el entrenamiento ANN. El Levenberg-Marquardt es un método de entrenamiento basado en híbridos que utiliza el método de descenso más pronunciado (descenso de gradiente) y el método de Gauss-Newton. Acelera la convergencia a una solución óptima y, por lo tanto, es eficaz para resolver problemas no lineales sobre los otros algoritmos de entrenamiento [15, 16]. El algoritmo introduce otro enfoque de aproximación a Hessian Matrix, que es similar al método de Gauss-Newton [15, 17]:

$$w_{ij+1} = w_{ij} + [J^T J + I]^{-1} J^T e_k \quad (6)$$

donde J denota la matriz jacobiana, e_k es el error en la red [17], w_{ij} es el peso actual y w_{ij+1} es el peso actualizado, \vec{y} es el factor de amortiguamiento.

Cuando \vec{y} es pequeño, el algoritmo de entrenamiento de Levenberg-Marquardt en la Ec. (6) se basa en el método de Gauss-Newton (matriz de confusión). Además, una característica de funcionamiento del receptor

y, por el contrario, se convierte en un algoritmo de descenso de gradiente cuando \vec{y} es grande. Por lo tanto, \vec{y} se ajusta en cada iteración para guiar el proceso de optimización y cambia entre esos dos algoritmos.

El número de neuronas en la capa de salida es el resultado y la predicción de la decisión del problema [4]. La capa de salida consta de colecciones de vector \vec{y} ción de CGPA previsto.

3.4 Criterios de evaluación del desempeño de la RNA

Para el análisis de datos posterior, con el fin de evaluar el rendimiento de ANN, este documento presenta varias perspectivas nuevas que mitigan la aparición de problemas de ajuste excesivo para garantizar la idoneidad del rendimiento de ANN. La evaluación incluye el cálculo del error cuadrático medio (MSE), análisis de regresión, histograma de error y matriz de confusión. Un modelo ANN bien entrenado debe tener un valor de MSE bajo (cerca de cero), lo que significa que los resultados previstos convergen estrechamente con los resultados

objetivo (t_{ij}). MSE se calcula como:

$$MSE = \frac{1}{\text{por } N_j} \sum_{j=1}^{N_j} \sum_{y=1}^{N_y} (y_{ij} - t_{ij})^2 \quad (7)$$

Como se mencionó, es posible un ajuste excesivo de la red entrenada si el valor de MSE obtenido es bajo. Esto indica además que ANN solo funciona bien en la etapa de entrenamiento, pero no en la fase de validación y prueba. Para mitigar esto, se realiza una regresión junto con el valor R calculado que demuestra la bondad del ajuste entre los resultados previstos y deseados [18]. El gráfico es útil para examinar el rendimiento del accesorio. Si se obtiene un ajuste deficiente (valor R bajo), se requieren más entrenamientos con la modificación de capas y neuronas ocultas.

Como se mencionó, otra forma de medir el rendimiento de ANN es tabular el histograma de error. El histograma de errores demuestra cómo se distribuyen los errores y la mayoría de los errores se producen cerca de cero. El error es simplemente la diferencia entre las salidas objetivo t_{ij} y las salidas previstas y_{ij} .

Para verificar el desempeño de ANN en términos de clasificaciones, se utiliza una matriz de confusión (también conocida como matriz de error). Una matriz de confusión de una clasificación binaria es una tabla de dos por dos que muestra valores de verdaderos negativos (TN), falsos negativos (FN), verdaderos positivos (TP) y falsos positivos (FP) resultantes de las clases de datos pronosticadas [19, 20]. La matriz de confusión permite medir tasas como la precisión de la predicción, la tasa de error, la sensibilidad, la especificidad y la precisión [19], que se incluyen en este documento (consulte [21] para obtener información sobre los cálculos de las tasas derivadas de la

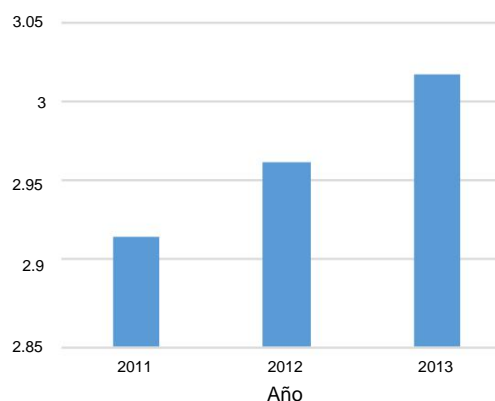


Fig. 1 Rendimiento promedio anual de CGPA de los estudiantes para el año de ingreso 2011–2013

También se emplea la curva (ROC) que detecta el intercambio entre la tasa de TP y la tasa de FP, y también el área bajo la curva ROC (AUC) [20]. ROC se usa ampliamente en el aprendizaje automático para verificar y evaluar el rendimiento de las clasificaciones [19].

4 resultados

Esta sección presenta los hallazgos estadísticos, la configuración de ANN y el desempeño basado en los datos educativos de rendimiento académico anual obtenidos de 1,000 estudiantes (175 mujeres y 810 hombres) para el año de entrada 2011–2013 de la Universidad Q.

4.1 Evaluaciones estadísticas

La Figura 1 presenta el desempeño del promedio anual de CGPA de los estudiantes para el año de ingreso 2011–2013. Además, el ANOVA para tres años académicos consecutivos con $p < 0,05$ indica que existe una diferencia significativa en el desempeño anual de CGPA de los estudiantes, con mejoras notables en su desempeño para los próximos dos años académicos.

Las pruebas t de dos muestras se realizan en tres tipos de datos de muestra para probar las hipótesis nulas sobre el género, la ubicación y los estudiantes repetidores de los estudiantes.

La Tabla 1 muestra el resultado de la prueba t del rendimiento de CGPA en función del género de los estudiantes. Dado que $p < 0,05$, hay diferencias significativas en el desempeño de CGPA para estudiantes masculinos y femeninos, donde en este caso las estudiantes obtienen mejores puntajes que los estudiantes masculinos. La Tabla 2 presenta otra prueba t del desempeño de CGPA de estudiantes basados en asentamientos urbanos y rurales. Los resultados de la prueba t con $p > 0,05$ concluyen que no existe una diferencia significativa en el rendimiento de CGPA en función de los diferentes tipos de liquidación. Finalmente, el

Tabla 1 Resultado de la prueba T de CGPA según el género de los estudiantes

| Género | Masculino | Mujer |
|----------------|-----------|-------|
| Frecuencia | 249 | 751 |
| CGPA medio | 2.54 | 3.11 |
| Diferencia | 0.22 | 0.18 |
| t-cal | y 17.33 | |
| t-grito | 1.65 | |
| d.f. | 391 | |
| p (una cola) p | 0.01 | |
| (dos colas) | 0.01 | |
| Decisión | Rechazar | |

Tabla 2 Resultado de la prueba T de CGPA basado en el tipo de asentamiento de los estudiantes

| basado en liquidación | Urbano | Rural |
|-----------------------|---------|-------|
| Frecuencia | 245 | 755 |
| CGPA medio | 2.93 | 2.98 |
| Diferencia | 0.28 | 0.24 |
| t-cal | y 1,32 | |
| t-grito | 1.64 | |
| d.f. | 390 | |
| p (una cola) p | 0.09 | |
| (dos colas) | 0.18 | |
| Decisión | Aceptar | |

Tabla 3 Resultado de la prueba T de estudiantes repetidores basados en CGPA

| Estudiantes repetidores | No | Sí |
|-------------------------|----------|------|
| Frecuencia | 711 | 289 |
| CGPA medio | 3.04 | 2.82 |
| Diferencia | 0.23 | 0.25 |
| t-cal | 6.04 | |
| t-grito | 1.64 | |
| d.f. | 513 | |
| p | 0.01 | |
| p (dos colas) | 0.01 | |
| Decisión | Rechazar | |

El resultado de la prueba t con $p < 0.05$ en la Tabla 3 indica que el desempeño promedio de CGPA de los estudiantes actuales y repetidores es estadísticamente significativo, con estudiantes actuales que obtienen mejores puntajes que los estudiantes repetidores.

La prueba de correlación se realiza para examinar el grado de relación más estrecha entre los resultados del examen de materias de ingreso y el resultado de CGPA. Los coeficientes de correlación entre cinco asignaturas del examen de ingreso y el desempeño del CGPA se muestran en la Tabla 4, en orden descendente: Inglés, Matemáticas, Chino, Ciencias Integrales y Prueba de Aptitud.

Se realizan pruebas ANOVA para examinar los efectos de la ocupación de los padres en el CGPA de los estudiantes resultantes. La ocupación de los padres se clasifica en cinco categorías y los efectos de la ocupación de la madre y el padre en el desempeño de los estudiantes se evalúan por separado, con los resultados correspondientes que se muestran en las Tablas 5 y 6.

La $p < 0,05$ en la Tabla 5 indica que las CGPA de los estudiantes son significativamente afectadas por la ocupación de las madres. Se puede inferir que uno de los grandes logros de los estudiantes se basa en motivaciones de su madre. Además, se puede explicar porque las madres suelen estar más preocupadas por el rendimiento académico de sus hijos, por lo que influyen mucho en el rendimiento de los estudiantes.

En contraste, no hay una diferencia significativa ($p > 0.05$) del CGPA promedio de los estudiantes basado en la ocupación laboral de los padres como se muestra en la Tabla 6, donde la motivación de los padres no es lo suficientemente fuerte para mejorar el desempeño académico entre los estudiantes. No es el alcance de este documento evaluar más a fondo la diferencia media de las unidades de trabajo individuales dentro de los padres y el CGPA resultante.

4.2 Ajustes de configuración de ANN

El modelado y las evaluaciones de ANN se realizan utilizando el software MathWorks MATLAB. La capa de entrada consta de 11 variables sobre los antecedentes de los estudiantes y los resultados de sus exámenes de ingreso, con la información de antecedentes

incluido el género, la ubicación, si los estudiantes repiten o no, el área escolar anterior, la ocupación de los padres y los resultados del examen de ingreso que incluye chino, inglés, matemáticas, ciencias integrales y prueba de competencia. La ANN modelada tiene dos capas ocultas, donde cada capa oculta consta de 30 neuronas, ya que tal configuración proporciona el mejor resultado a lo largo de varias simulaciones con diferentes configuraciones de capas ocultas y neuronas. Cada capa oculta con 30 neuronas se alimenta a una sola neurona de salida que lleva la decisión de la variable, que es la predicción del CGPA de los estudiantes.

Se utiliza la función de activación de la tangente hiperbólica. Se deben tomar decisiones para dividir el conjunto de datos en proporción de entrenamiento, validación y prueba. Se mezclan aleatoriamente muestras de datos de 1.000 estudiantes y 0,7 de las muestras mixtas se utilizan para formación, 0,15 para validación y el 0,15 restante para pruebas. Siguiendo [14], PCA se mantiene solo con suficientes componentes para explicar el 95% de la varianza.

Durante la fase de entrenamiento y aprendizaje, el algoritmo Levenberg-Marquardt en Eq. (6) se usa para determinar los pesos óptimos que se alimentan a la siguiente capa de entrada. El factor de amortiguamiento γ se establece en 0.001. La época de entrenamiento se establece en 1500. El entrenamiento de ANN se realiza de forma continua y finaliza cuando el error de validación no disminuye durante seis iteraciones durante el proceso de validación.

Por lo general, la validación protege el sobreentrenamiento de ANN.

Tabla 4 Coeficiente de correlación de cinco áreas temáticas y CGPA

| | Chino | Matemáticas | Inglés | Ciencias integrales | Prueba de competencia CGPA | |
|----------------------|-------|-------------|--------|---------------------|----------------------------|---|
| chino | 1 | | | | | |
| Matemáticas | 0.10 | 1 | | | | |
| inglés | 0.30 | 0.21 | 1 | | | |
| ciencia comprensiva | 0.33 | 0.45 | 0.54 | 1 | | |
| Test de proficiencia | 0.25 | 0.22 | 0.46 | 0.32 | 1 | |
| CGPA | 0.49 | 0.58 | 0.60 | 0.39 | 0.32 | 1 |

Tabla 5 ANOVA de una vía contra las ocupaciones de las madres

| Fuente | suma de cuadrados | d.f. | Cuadrado medio | F-cal | p -value | F-crítico |
|------------------|-------------------|------|----------------|-------|------------|-----------|
| Entre grupos | 2.93 | 5 | 0.59 | 2.39 | 0.04 | 2.22 |
| Dentro de grupos | 243.89 | 995 | 0.25 | | | |
| Total | 246.82 | 1000 | | | | |

Tabla 6 ANOVA de una vía contra las ocupaciones de los padres

| Fuente | suma de cuadrados | d.f. | Cuadrado medio | F-cal | p -value | F-crítico |
|------------------|-------------------|------|----------------|-------|------------|-----------|
| Entre grupos | 1.38 | 5 | 0.28 | 1.12 | 0.35 | 2.22 |
| Dentro de grupos | 245.58 | 995 | 0.25 | | | |
| Total | 246.96 | 1000 | | | | |

4.3 Resultados de la simulación ANN

Después del entrenamiento ANN, PCA puede mantener cinco componentes, y la varianza explicada por componente (en orden) es: 62,9 %, 16,6 %, 8,8 %, 5,1 %, 2,8 %, 2,2 %, 0,9 %, 0,6 %, 0,1 %, 0 % (donde las variaciones de los componentes menos importantes se ocultan de la visualización). El rendimiento de ANN con MSE \bar{y} 0,27 o 6,9 % (< 10 %) indica un rendimiento suficientemente bueno de ANN para completar las carreras de entrenamiento y validación.

Los gráficos de regresión del rendimiento de ANN se muestran en la Fig. 2. Para un buen ajuste de datos, los datos deben tener las salidas previstas y_{ij} muy cerca de las salidas objetivo t_{ij} . En

En este caso, los gráficos de regresión resultantes logran fts marginalmente buenos con un valor R general de 0,64.

El histograma de errores evalúa las distribuciones de errores en función de las predicciones ANN resultantes. La distribución de errores se muestra en la Fig. 3. La mayoría de los errores ocurrieron cerca del punto cero (eje horizontal), y los errores disminuyen gradualmente al alejarse del punto cero. Esto prueba que ANN realiza la predicción con éxito con distribuciones de error aceptables, como se muestra en la Fig. 3.

La Figura 4 presenta los resultados de la matriz de confusión para un clasificador binario resultante de dos clases (género de los estudiantes). Como puede verse, 136 de 249 estudiantes varones y 712 de 751 estudiantes mujeres han sido clasificados correctamente.

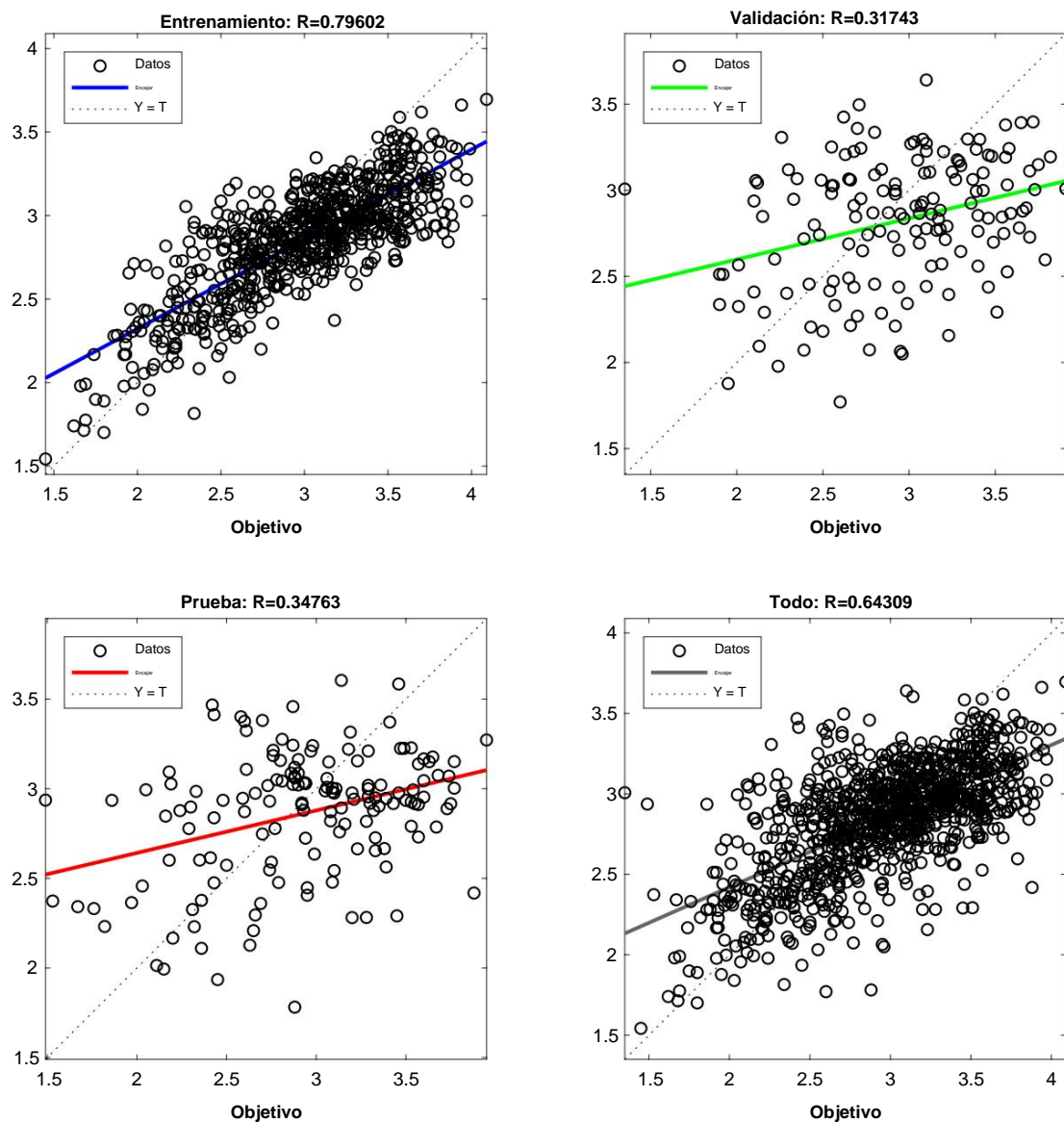


Fig. 2 Gráfico de regresión para el modelo ANN: arriba a la izquierda: datos de entrenamiento; abajo a la izquierda: datos de prueba; Arriba a la derecha: datos de validación; abajo a la derecha: regresiones resultantes

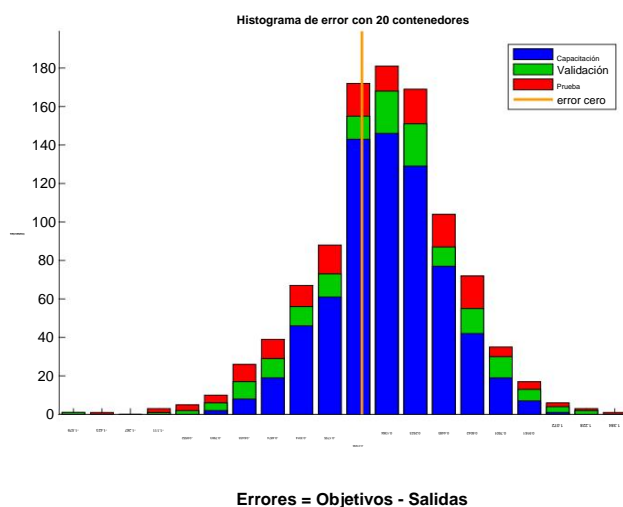


Fig. 3 Gráfico de histograma de error

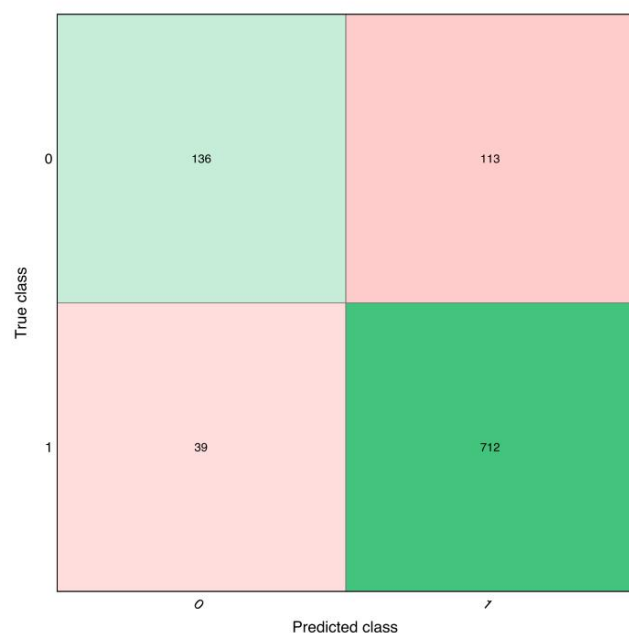


Fig. 4 Clasificador binario de matriz de confusión basado en resultados de clasificación por ANN

De manera similar, la figura 5 muestra las distribuciones porcentuales de la matriz de confusión de la figura 4. Las tasas de verdadero positivo (TP) se calculan como 55 % (hombre) y 95 % (mujer) y, a su vez, las tasas de falso negativo (FN) se computan como 45% (masculino) y 5% (femenino). Estas altas tasas de FN para la clasificación de los estudiantes varones pueden deberse a la alta proporción desequilibrada del género de los estudiantes (249 estudiantes varones frente a 751 estudiantes mujeres).

El uso de la matriz de confusión permite la determinación de tasas tales como exactitud, tasa de error, sensibilidad, especificidad y precisión, y los resultados son los siguientes (en porcentaje):

Precisión = 84,8%

Error = 15,2%

Sensibilidad = 94,8%

Especificidad = 54,6%

Precisión = 86,3%

La curva ROC de ANN se muestra en la Fig. 6. El área bajo la curva ROC (AUC) se usa comúnmente para evaluar la efectividad de la precisión de ANN en la predicción y clasificación, donde un AUC de 1 representa una prueba perfecta. Según el gráfico, se logra el valor AUC de 0,86, con una tasa de intercambio de TP y FP de 0,05 y 0,55. Esto demuestra el éxito suficiente del estudio ANN en la predicción y clasificación del rendimiento de los estudiantes.

5. Conclusión

En este artículo, ANN se utiliza para evaluar y predecir el CGPA de los estudiantes utilizando los datos sobre su entorno socioeconómico y los resultados del examen de ingreso de los estudiantes de pregrado de una universidad china. La etapa inicial consiste en análisis estadísticos de la información de antecedentes de los estudiantes que han indicado una mejora significativa en el caso del CGPA anual de los estudiantes; las alumnas obtienen puntuaciones significativamente mejores que los alumnos varones; no existe una diferencia significativa en el desempeño de los estudiantes de las zonas rurales y urbanas; y los alumnos que no repiten se desempeñan significativamente mejor que los alumnos que repiten.

El análisis de correlación indica que el resultado del examen de inglés en el examen de ingreso tiene la correlación más fuerte con el CGPA del estudiante. La madre de los estudiantes juega un papel significativamente más importante en el desempeño académico de los estudiantes que el padre.

Para evaluar el rendimiento de ANN, se introducen cálculos de error cuadrático medio (MSE), análisis de regresión, histograma de error y matriz de confusión para garantizar la idoneidad del rendimiento de ANN para mitigar la aparición de problemas de sobreprecio. En general, la ANN ha logrado una buena precisión de predicción del 84,8 % y un buen valor de AUC de 0,86. El buen desempeño de ANN hace que ANN sea una técnica tan útil para fines de evaluación educativa. Sin embargo, ANN tiene un desempeño deficiente en las clasificaciones de los estudiantes según su género, ya que se obtienen como resultados altas tasas de Falsos Negativos (FN). Esto probablemente se deba a la alta proporción de desequilibrio de dos tipos diferentes de datos de muestra (sexo de los estudiantes) aplicados. Estas deficiencias se pueden mitigar mediante la introducción de nuevas muestras con un tipo de población de muestra más equilibrado.

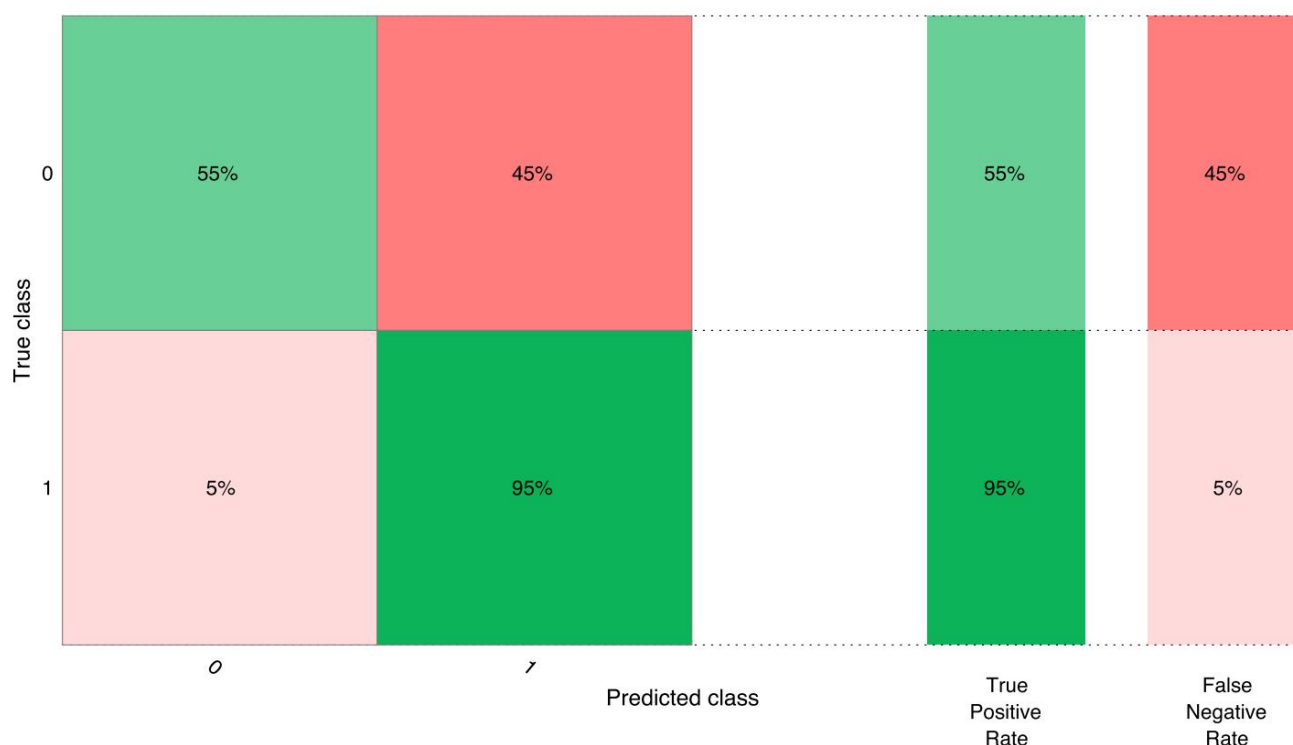


Fig. 5 Clasificador binario de matriz de confusión basado en resultados de clasificación por ANN en términos de porcentaje

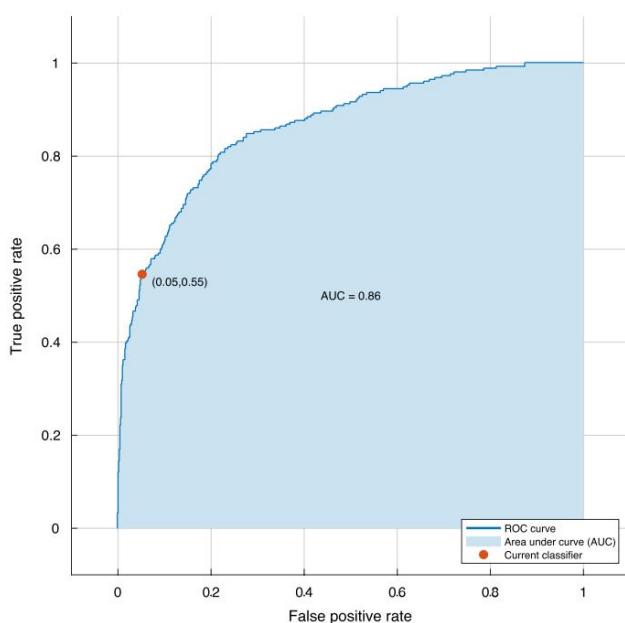


Fig. 6 Curva ROC de ANN. Área AUC bajo la curva ROC

A la luz de los resultados, se cree firmemente que los entornos de modelado educativo que utilizan ANN pueden ayudar a evaluar el rendimiento de los estudiantes. Los futuros modelos pueden incluir otros atributos como el papel del profesorado,

las clases curriculares adicionales, la retroalimentación del curso y el sistema de aprendizaje electrónico son tareas inmediatas para complementar un modelo completo de ANN, mientras se enfoca en la precisión de predicción de ANN. Creemos firmemente que la educación futura con un entorno analítico basado en inteligencia artificial puede mejorar aún más la capacidad de los estudiantes para lograr un alto rendimiento académico dentro de las instituciones académicas.

Financiación Esta investigación está parcialmente financiada por el Proyecto DIA170362 (2017-2020) en el marco del 13.º Plan Quinquenal de Proyectos Nacionales de Ciencias de la Educación 2017.

Cumplimiento de normas éticas

Conflicto de intereses Los autores declaran no tener ningún conflicto de intereses.

Acceso abierto Este artículo se distribuye bajo los términos de Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), que permite el uso, la distribución y la reproducción sin restricciones en cualquier medio, siempre que otorgue el crédito correspondiente al autor o autores originales y la fuente, proporcione un enlace a la licencia Creative Commons e indique si se realizaron cambios.

Referencias

1. Tu JV (1996) Ventajas y desventajas del uso de redes neuronales artificiales versus regresión logística para predecir resultados médicos. *J Clin Epidemiol* 49:1225–1231
2. Kardan AA, Sadeghi H, Ghidary SS, Sani MR (2013) Predicción de la selección de cursos de los estudiantes en institutos de educación superior en línea utilizando redes neuronales. *Computación Educativa* 65:1–11
3. Hu B (2017) Investigación de evaluación de la calidad docente basada en redes neuronales para la educación física universitaria. En: *IEEE*, págs. 290–293. <https://doi.org/10.1109/ICSGEA.2017.155>
4. Özçelik S, Hardalaç N (2011) Las mediciones estadísticas y el análisis de redes neuronales del efecto de la educación musical en la audición y la detección musicales. *Aplicación de sistema experto* 38:9517–9521
5. Vandamme JP, Meskens N, Superby JF (2007) Predicción del rendimiento académico mediante métodos de minería de datos. *Educ Econ* 15(4):405–419
6. Kose U, Arslan A (2017) Optimización del autoaprendizaje en el curso de ingeniería informática: un sistema de software inteligente respaldado por una red neuronal artificial y un algoritmo de optimización de vórtice. *Comput Appl Eng Educ* 25: 142–156
7. Rashid TA, Ahmad HA (2016) Uso de red neuronal con optimización de enjambre de partículas. *Comput Appl Eng Educ* 24: 629–638
8. Villasenor EA, Arencibia-Jorge R, Carrillo-Calvet H (2017) Caracterización multiparamétrica de perfiles de desempeño cienciométrico asistidos por redes neuronales: un estudio de instituciones de educación superior mexicanas. *Cienciometría* 110:77–104
9. Mustafa HMH, Al-Hamadi A, Al-Ghamdi SA, Hassan MM, Khedr AA (2013) Sobre la evaluación del rendimiento académico de los estudiantes considerando las diferencias individuales categorizadas en la educación en ingeniería (enfoque de redes neuronales). En: *IEEE*, págs. 1–10. <https://doi.org/10.1109/ITHET.2013.6671003>
10. Bajaj R, Sharma V (2018) Educación inteligente con inteligencia artificial basada en la determinación de estilos de aprendizaje. *Proc Comput Sci* 132: 834–842
11. Moridis CN, Economides AA (2009) Predicción del estado de ánimo de los estudiantes durante una prueba en línea utilizando un método basado en fórmulas y redes neuronales. *Educación informática* 53: 644–652
12. Bahadýr E (2016) Uso de redes neuronales y análisis de regresión logística para predecir el éxito académico de los futuros profesores de matemáticas al ingresar a la educación de posgrado. *Educ Sci Teoría Práctica* 16(3):943–964
13. Zhang Q, Kuldip C, Devabhaktuni VK (2003) Red neuronal artificial para el diseño de rf y microondas: de la teoría a la práctica. *IEEE Trans Microw Theory Tech* 51 (4): 1339–1350
14. MathWorks (2013) Selección de funciones y transformación de funciones mediante la aplicación de aprendizaje de clasificación. <https://uk.mathworks.com/help/stats/feature-selection-and-feature-transformation.html#buwh6hc-1>. Consultado el 17 de junio de 2019
15. Yu H, Wilamowski BM (2011) *Leverberg Marquardt manual de electrónica industrial de formación, sistemas inteligentes*, vol 5, 2.ª ed. CRC Press, Boca Ratón
16. Wilson P, Mantooth HA (2013) *Ingeniería basada en modelos para sistemas electrónicos complejos*. Newness, Oxford
17. MathWorks (2019a) trainlm–Leverberg–Marquardt backpropagation. <http://uk.mathworks.com/help/nnet/ref/trainlm.html>. Consultado el 17 de junio de 2019
18. MathWorks (2019b) Estimación de la grasa corporal. <https://uk.mathworks.com/help/nnet/examples/body-fat-estimation.html>. Consultado el 17 de julio de 2019
19. Saito T, Rehmsmeier M (2015) La gráfica de recuperación de precisión es más informativa que la gráfica ROC cuando se evalúan clasificadores binarios en conjuntos de datos desequilibrados. *PLoS ONE* 10(3):1–21
20. Grossman MR, Cormack GV (2013) Edrm page y el glosario de grossman cormack de revisión asistida por tecnología, con prólogo de john m. facciola, juez magistrado estadounidense. *Ley de Tribunales Federales Rev* 7(1):1–33
21. Data School (2014) Guía simple de terminología de matriz de confusión. <https://www.dataschool.io/simple-guide-to-confusion-matrix-terminology/>. Consultado el 7 de julio de 2018

Nota del editor Springer Nature se mantiene neutral con respecto a los reclamos jurisdiccionales en los mapas publicados y las afiliaciones institucionales.