

# Comparación del rendimiento de la regresión múltiple y la red neuronal artificial en la determinación del orden de importancia de los predictores en la investigación educativa<sup>1,2</sup>

## Comparison of multiple regression and artificial neural network performances in determining the order of importance of predictors in educational research

<https://doi.org/10.4438/1988-592X-RE-2023-399-568>

**Emre Toprak**

<https://orcid.org/0000-0002-4131-4888>

*Erciyes University*

**Ömür Kaya Kalkan**

<https://orcid.org/0000-0001-7088-4268>

*Pamukkale University*

### Resumen

Los estudios que tienen como objetivo determinar los rangos de importancia de una o más variables predictoras en la variable predicha se encuentran con frecuencia en la educación. La regresión múltiple (RM) y la red neuronal artificial (RNA) son ampliamente utilizadas en este tipo de investigación. El presente estudio compara el desempeño del rango de importancia predictiva de los métodos RM y RNA. Para este propósito, dos conjuntos de datos reales separados, en los que se cumplen los supuestos de RM y las variables predictoras

---

<sup>1</sup> Los autores declaran que una parte de este estudio se presentó como una presentación de resumen oral en el 2do Congreso ICES celebrado del 18 al 19 de junio de 2019 en Turquía y 6to Congreso EJER, celebrado del 19 al 22 de junio de 2019 en Turquía.

<sup>2</sup> Los autores desean expresar su más profundo agradecimiento a Hakan DEMİRTAS por su contribución.

son *continuas* o *discretos*, y se utilizaron datos de simulación generados al considerar las relaciones en estos conjuntos de datos. Se utilizaron sesgos relativos absolutos (SRA) y errores cuadráticos medios (ECM) para comparar el rendimiento de los métodos. Los resultados de la investigación mostraron que el aumento en el tamaño de la muestra tuvo un efecto de mejora en los SRAs y ECMs de los métodos. Además, si los predictores son *continuas*, se puede recomendar a los investigadores que elijan RM o RNA. Sin embargo, en los casos en que los predictores sean *discretos* y el número de predictores sea tres o más, se recomienda el uso de RNA. Para obtener estimaciones óptimas con ambos métodos, se recomienda que los investigadores utilicen un tamaño de muestra de al menos 200.

*Palabras clave:* análisis de regresión múltiple, redes neuronales artificiales, predictor continua, predictor discreto, orden de importancia, investigación correlacional predictiva.

### **Abstract**

Studies aiming to determine the importance rankings of one or more predictor variables on the predicted variable are frequently encountered in education. Multiple Regression (MR) and artificial neural network (ANN) are widely used in this type of research. The present study compares the predictive importance rank performances of MR and ANN methods. For this purpose, two separate real data sets, in which MR assumptions are met and the predictor variables are *continuous* or *discrete*, and simulation data generated by considering the relationships in these data sets were used. Absolute relative bias (ARB) and mean square errors (MSE) were used to compare the methods' performances. The results of the research showed that the increase in sample size had an improving effect on the ARBs and MSEs of the methods. In addition, if the predictors are *continuous*, researchers may be advised to choose either MR or ANN. However, in cases where the predictors are *discrete* and the number of predictors is three or more, the use of ANN is recommended. In order to obtain optimal estimations with both methods, it is recommended that researchers use a sample size of at least 200.

*Keywords:* multiple regression analysis, artificial neural networks, continuous predictor, discrete predictor, order of importance, predictive correlational research.

## Introducción

El tipo predictivo de estudios correlacionales se encuentra con frecuencia en los campos de la educación. En los estudios correlacionales de tipo predictivo, si existe una correlación adecuada entre dos variables, la puntuación de una variable puede utilizarse para predecir la puntuación de la otra variable (Fraenkel et al., 2011). En estos estudios, la variable que se estima puede expresarse como la variable predicha (dependiente/criterio), y la/s variable/s utilizada/s para la estimación puede/n expresarse como la/s variable/s predictiva/s (independiente/s).

En estudios correlacionales de tipo predictivo realizados en los campos de la educación variables relacionadas con las características cognitivas (e.g., Akgül, 2019; Aramburo et al., 2017; Lee, 2016, etc.), características afectivas o psicomotoras (e.g., Arthur & Doverspike, 2001; Atabey, 2020; Ivcevic & Eggers, 2021; Marroquin & Nolen-Hoeksema, 2015, etc.) se abordan. En estos estudios se puede determinar el orden de importancia de las variables predictoras sobre la variable predicha. Además, se pueden revelar resultados importantes que revelan las relaciones entre las estructuras latentes (por ejemplo, logros, motivación, depresión, etc.) que no se pueden observar directamente pueden ser revelados. De este modo, se puede proporcionar una base científica a los investigadores de la educación, a los profesionales de la educación y a los responsables de la toma de decisiones en relación con las cuestiones que deben priorizar y/o mejorar.

Cuando se examinan los estudios correlacionales predictivos realizados en los campos de la educación, se observa que los análisis de regresión lineal simple y múltiple se utilizan principalmente para predecir la variable dependiente (e.g., Akhtar & Herwig, 2019; Alandete, 2015; Bergold & Steinmayr, 2018; Garcini et al., 2013; Teressa & Bekele, 2021, etc.). Sin embargo, los análisis de regresión tienen muchas suposiciones que deben cumplirse, y puede haber situaciones en las que no se cumplan todas estas suposiciones. Esta situación pone de manifiesto la necesidad de emplear diferentes métodos en los estudios de predicción. Uno de los métodos que se pueden utilizar como alternativa al análisis de regresión múltiple (RM) en los estudios de predicción son las redes neuronales artificiales (RNA). RNA se ha utilizado en ciencias sociales, ingeniería, negocios y muchos otros campos (e.g., Bayru, 2007; Cansız et al., 2020; Okkan & Mollamahmutoğlu, 2010; Tolon, 2007, etc.)

durante mucho tiempo, recientemente ha sido utilizado con frecuencia en la investigación educativa (e.g., Kalkan & Coşguner, 2021; Moreetsi & Mbako, 2008; Toprak & Gelbal, 2020, etc.). Los métodos RM y RNA utilizados en la investigación se presentan brevemente a continuación.

## Regresión múltiple (RM)

Los análisis de regresión son un conjunto de técnicas estadísticas que permiten evaluar la relación entre una variable dependiente y una o más variables independientes (Tabachnick & Fidell, 2013). RM se ocupa de predecir una variable dependiente en función de un conjunto de estimadores (Stevens, 2009). Además, RM permite determinar el total de la variación explicada por las variables predictoras sobre la variable dependiente, examinar la significancia estadística de las variables predictoras e interpretar la relación entre las variables predictoras y la variable dependiente (Büyüköztürk, 2014). Además, la RM es una técnica bastante eficiente para analizar los efectos totales y separados de las variables predictoras sobre una variable dependiente (Pedhazur, 1997).

En RM, las variables predictivas deben ser continuas variables al menos en una escala de intervalo; sin embargo, *discretos* las variables también pueden incluirse en el análisis después de que se hayan definido como variables ficticias (Büyüköztürk, 2014). Además, RM tiene algunos supuestos que cumplir, como normalidad, linealidad, homocedasticidad de residuos, independencia de errores, multicolinealidad, tamaño de muestra adecuado y examen de valores atípicos (Tabachnick & Fidell, 2013).

## Redes neuronales artificiales (RNA)

El comienzo de los estudios RNA, que se basa en el principio de funcionamiento del cerebro humano, se remonta a la década de 1940. Las RNA también se expresan como programas informáticos que imitan redes neuronales biológicas en el cerebro humano, en forma de simulación del sistema nervioso biológico (Elmas, 2003). Una célula nerviosa biológica consta de cuatro componentes esenciales: dendrita, núcleo, axón y sinapsis (Öztemel, 2012). De manera similar, una neurona

artificial consta de las entradas creadas por la información que proviene del mundo exterior, los pesos por los cuales se determina el nivel de importancia de la información, la función de suma que proporciona el cálculo de la información neta que ingresa a la célula y la función de activación que determina la salida a producir (Baş, 2006; Öztemel, 2012).

Algunas ventajas son efectivas en el uso generalizado de RNA. Se pueden enumerar como ser capaz de resolver problemas complejos, adaptarse a nuevas situaciones con pequeños cambios, generalizar en base a los resultados obtenidos, ser más rápido que los métodos estadísticos tradicionales, trabajar con datos incompletos, manejar datos defectuosos o multidimensionales con tolerancia a los errores, y el reconocimiento de patrones en datos perdidos (Çırak & Çokluk, 2013; Haykin, 1994; Öztemel, 2012; Simpson, 1990). Estas ventajas de RNA lo hacen preferible a los métodos basados en regresión.

Hay estudios en la literatura realizados con el fin de comparar los rendimientos de predicción de RM y RNA. En estos estudios, se examinaron los rendimientos de los métodos en diferentes parámetros, como los cuadrados medios de error, los porcentajes de predicción, los coeficientes de determinación y los diagramas de dispersión (e.g., Akbilgiç & Kesintürk, 2008; Cansız et al., 2020; Gorr et al., 1994; Lykourentzou et al., 2009; Okkan & Mollamahmutoglu, 2010; Turhan et al., 2013; Zaidah & Daliela, 2007, etc.). Hasta donde sabemos, no se ha encontrado ninguna investigación que examine el rendimiento de ambos métodos para determinar el orden de importancia de las variables predictoras en conjuntos de datos reales y de simulación según el tipo de predictor, el número de predictores y el tamaño de la muestra. Por tanto, los resultados de la investigación pueden contribuir a la determinación de las limitaciones de los métodos en la práctica; además, se piensa que puede proporcionar inferencias para investigadores y profesionales. Por esa razón, el presente estudio tiene como objetivo comparar los rangos de prioridad obtenidos de RM y RNA, para determinar el orden de importancia de las variables sobre conjuntos de datos reales y de simulación de diferentes tamaños de muestra que consisten en variables *continuas* o variables *discretos*. Para ello, se buscaron respuesta a las siguientes preguntas de investigación (PI):

**PI1.** ¿Cuáles son las relaciones entre las variables en conjuntos de datos reales que consisten en variables predictivas *continuas* o *discretos*?

**PI2.** Cuando las variables predictoras son *continuas*;

- a. ¿Cuál es el orden de importancia obtenido por RM y RNA del conjunto de datos real?
- b. ¿De qué manera el orden de importancia, el sesgo relativo absoluto (SRA) y el error cuadrático medio (ECM) obtenidos por RM y RNA a partir de conjuntos de datos de simulación con diferentes tamaños de muestra (100, 200, 400, 800) y el número de predictores (2, 3, 4) cambió?

**PI3.** Cuando las variables predictoras son *discretos*;

- a. ¿Cuál es el orden de importancia obtenido por RM y RNA del conjunto de datos real?
- b. ¿De qué manera el orden de importancia, el sesgo relativo absoluto (SRA) y el error cuadrático medio (ECM) obtenidos por RM y RNA a partir de conjuntos de datos de simulación con diferentes tamaños de muestra (100, 200, 400, 800) y el número de predictores (2, 3, 4) cambió?

## Método

### Grupo de estudio

En el ámbito de la investigación, se utilizaron dos conjuntos de datos reales diferentes pertenecientes a los estudios de Toprak y Kalkan (2019a, 2019b) para comparar el rendimiento de los métodos RM y RNA en la determinación del orden de importancia de las variables *continuas* y *discretos*. El primer conjunto de datos consta de las respuestas dadas por 445 estudiantes (324 mujeres, 121 hombres) matriculados en tres escuelas secundarias diferentes en Turquía (Toprak & Kalkan, 2019a). Este conjunto de datos se utilizó para comparar el orden de importancia de los métodos cuando las variables predictoras eran *continuas*.

El segundo conjunto de datos consta de las respuestas dadas por 992 estudiantes (540 mujeres, 452 hombres) matriculados en cinco escuelas secundarias diferentes en Turquía (Toprak & Kalkan, 2019b). Este conjunto de datos se usó para comparar el orden de los métodos cuando las variables predictoras eran *discretos*.

## **Instrumentos de recolección de datos**

El primer conjunto de datos del presente estudio consta de las puntuaciones de la Escala de flexibilidad cognitiva, la Escala de estrategias de afrontamiento y la Escala de expectativa de competencia utilizadas por Toprak y Kalkan (2019a) en su investigación. Estas escalas se presentan a continuación, respectivamente.

### **Escala de flexibilidad cognitiva**

La escala desarrollada por Martin y Rubin (1995) y adaptada al turco por Çelikkaleli (2014) consta de 12 ítems. El coeficiente de fiabilidad de la escala Alfa de Cronbach ( $Cr \alpha$ ) fue de .71. La escala tiene una estructura factorial unidimensional.

### **Escala de estrategias de afrontamiento**

La escala, desarrollada por Spirito et al. (1988) y adaptado al turco por Bedel et al. (2014), consta de 11 ítems y tres subdimensiones: afrontamiento activo, afrontamiento evitativo y afrontamiento negativo. El coeficiente de fiabilidad de  $Cr \alpha$  de la escala fue de .72 para afrontamiento activo, .70 para afrontamiento evitativo, y .65 para afrontamiento negativo.

### **Escala de expectativa de competencia**

La escala, desarrollada por Muris (2001) y adaptada al turco por Çelikkaleli et al. (2006), consta de 23 ítems y tres subdimensiones: expectativa de competencia académica, expectativa de competencia social y expectativa de competencia emocional. El coeficiente de confiabilidad  $Cr \alpha$  se informó como .64 para la expectativa de competencia académica, .69 para la expectativa de competencia emocional, y .71 para la expectativa de competencia social.

El segundo conjunto de datos de la presente investigación consiste en los datos obtenidos por Toprak y Kalkan (2019b) con el formulario de información personal en sus estudios. Este formulario de información y las variables incluidas en el estudio se presentan brevemente a continuación.

## Formulario de datos personales

Este formulario consta de variables *discretos* determinadas como predictoras de los puntajes de rendimiento del curso de matemáticas, como el género, el nivel de ingresos familiares, el nivel de educación de la madre y el nivel de educación del padre (Tabla I).

**TABLA I.** Estadísticas descriptivas sobre los predictores de las puntuaciones de rendimiento en los cursos de matemáticas

Variables		f	%
Género	Femenino	540	54.4
	Masculino	452	45.6
Nivel de ingresos familiares	2000 lira turca (TL) y menos	220	22.2
	2001-4000 TL	414	41.7
	4001-6000 TL	197	19.9
	6001 TL y superior	161	16.2
Nivel educativo de la madre	Hasta escuela primaria	280	28.2
	Escuela secundaria	259	26.1
	Liceo	279	28.1
	No licenciado y superior	174	17.5
Nivel educativo del padre	Hasta escuela primaria	166	16.7
	Escuela secundaria	247	24.9
	Liceo	290	29.2
	No licenciado y superior	289	29.1

## Análisis de los datos

Se utilizaron los métodos RM y RNA para determinar el orden de importancia de las variables. Para este propósito, se utilizaron dos conjuntos de datos reales diferentes y conjuntos de datos de simulación generados al considerar estos conjuntos de datos. Los datos faltantes, la adecuación del tamaño de la muestra y el análisis de valores atípicos se realizaron para evaluar el ajuste de los conjuntos de datos reales para RM, y se probaron la normalidad, la linealidad, la homocedasticidad



de los residuos, la independencia de los errores y los supuestos de multicolinealidad.

Las distancias de Mahalanobis obtenidas en la determinación de los valores atípicos se compararon con el valor crítico de chi-cuadrado ( $\alpha = .001$ ). Al evaluar la adecuación del tamaño de la muestra, se tuvo en cuenta el número de predictores  $x$  15 propuesto por Stevens (2009) y  $50 + 8 x$  el número de predictores propuesto por Tabachnick y Fidell (2013). Se utilizaron gráficos de probabilidad normal (P-P) y gráficos de dispersión de residuos estandarizados de regresión para controlar los supuestos de normalidad, linealidad y covarianza de los residuos. Para determinar el problema de multicolinealidad, se consideraron los valores de la correlación bivariada, la tolerancia, el índice de condición (IC) y el factor de inflación de la varianza (FIV). Se utilizaron datos reales y datos de simulación en los análisis RM y RNA.

## Datos reales

Se utilizaron dos conjuntos de datos diferentes para el análisis de datos reales. En el conjunto de datos1, todas las variables son *continuas*, y la variable dependiente es la flexibilidad cognitiva. Las expectativas de competencia social, las expectativas de competencia académica, las expectativas de competencia emocional y las variables de afrontamiento activo se utilizaron para predecir la flexibilidad cognitiva (Toprak & Kalkan, 2019a).

En el conjunto de datos2, la variable dependiente son los puntajes de rendimiento del curso de matemáticas (variable *continua*). Variables *discretos* de género, ingreso familiar, nivel educativo del padre y nivel educativo de la madre se utilizaron como predictores de los puntajes de rendimiento del curso de matemáticas (Toprak & Kalkan, 2019b).

## Diseño de simulación

Se utilizó el software R (Core team, 2017) para generar datos de simulación. Se crearon 24 ( $2 \times 4 \times 3$ ) condiciones de estudio cruzando 2 tipos de variables (*continuas*, *discretos*), 4 tamaños de muestra (100,

200, 400, 800), y se generaron 3 números de predictores (2, 3, 4) y cien conjuntos de datos para cada condición de estudio.

Usando la varianza, la covarianza y los valores medios obtenidos del conjunto de datos real (conjunto de datos 1) que consta de variables *continuas* se generaron datos de simulación. Se utilizó el paquete SimDesign (Chalmers & Adkins, 2020) para generar estos conjuntos de datos.

Utilizando los valores de correlación, media y desviación estándar obtenidos del conjunto de datos real (conjunto de datos 2) que consta de predictores *discretos* se generaron datos de simulación. Se utilizó el paquete OrdNor (Amatya & Demirtas, 2015) para generar estos datos. Los valores de la variable dependiente (puntajes de rendimiento del curso de matemáticas) en los conjuntos de datos generados se redondearon al número entero más cercano y se limitaron al rango de 0 a 100 para ajustarse a los valores reales.

Los predictores *discretos* se codificaron como variables ficticias y se incluyeron en el análisis. La primera categoría de cada variable se determinó como grupo de referencia (Tabla I). Para determinar el orden de importancia de las variables *discretos* se tuvo en cuenta la media de los coeficientes estandarizados obtenidos de todas las categorías de la variable relevante excepto el grupo de referencia.

Para comparar el desempeño de los métodos se utilizó el orden de importancia de las variables, SRA y ECM. Los valores de los parámetros obtenidos de conjuntos de datos reales se asumieron como valores verdaderos y el SRA se calculó con estimaciones obtenidas de datos de simulación. Para el cálculo de la SRA se utilizó la siguiente fórmula (Bandalos & Leite, 2013).

$$SRA(\hat{\theta}_i) = \sum_{j=1}^{n_r} \left( \frac{|\hat{\theta}_{ij} - \theta_i|}{\theta_i} \right) / n_r \quad (1)$$

Se utilizó la fórmula del error cuadrático medio (ECM) para determinar la eficiencia de las estimaciones de parámetros (Bandalos & Leite, 2013).

$$ECM(\hat{\theta}_i) = \sum_{j=1}^{n_r} \frac{(\hat{\theta}_{ij} - \theta_i)^2}{n_r} \quad (2)$$

En las Ecuaciones 1 y 2, es  $j$  la estimación de muestra del  $i$  el verdadero valor del parámetro es el número de replicación. SRA se puede interpretar como porcentaje de SRA multiplicándolo por 100.

Para el análisis RM se utilizó la R y para el análisis RNA se utilizó el programa estadístico para ciencias sociales (SPSS; IBM, 2020). El Modelo de Perceptrón Multicapa se utilizó en el análisis RNA; El 70% de los conjuntos de datos se eligieron como muestras de entrenamiento y el 30% como muestra de prueba. La función tangente hiperbólica se usó para las celdas en la capa oculta y la función softmax se usó para las celdas en la capa de salida. En cada análisis realizado con RNA se utilizan diferentes datos de entrenamiento y prueba. Por esta razón, cada conjunto de datos en la condición relevante se analizó 100 veces para garantizar la estabilidad.

## Resultados

La idoneidad de los supuestos de tamaño de la muestra, valores atípicos, normalidad, linealidad, homocedasticidad de los residuos y multicolinealidad (por ejemplo, tolerancia, índice de condición, factor de inflación de la varianza, etc.) se examinaron antes de realizar los análisis de RM y se vio que se cumplieron estos supuestos. Luego, se presentan los resultados relacionados con los problemas de investigación respectivamente.

### Correlación entre variables en conjuntos de datos reales que consisten en variables predictoras continuas o discretos

Las correlaciones bivariadas entre las variables en los conjuntos de datos reales se presentan en la Tabla II, respectivamente.

**TABLA II.** Correlaciones en el conjunto de datos reales que consta de variables *continuas* y *discretos*

<b>Variabales continuas</b>	<b>FC</b>	<b>ECS</b>	<b>ECA</b>	<b>ECE</b>	<b>AA</b>
FC	–				
ECS	.439**	–			
ECA	.424**	.306**	–		
ECE	.425**	.418**	.338**	–	
AA	.401**	.291**	.379**	.483**	–
<b>Variabales discretos</b>	<b>PRCM</b>	<b>G</b>	<b>IF</b>	<b>NEP</b>	<b>NEM</b>
PRCM	–				
G	-.201**	–			
IF	.307**	.016	–		
NEP	.320**	-.067**	.425*	–	
NEM	.324**	-.031**	.457**	.589**	–

\*\*p<.01; \*p<.05. Nota. FC: flexibilidad cognitiva; ECS: expectativa de competencia social; ECA: expectativa de competencia académica; ECE: expectativa de competencia emocional; AA: afrontamiento activo; PRCM: puntajes de rendimiento del curso de matemáticas; G: género; IF: ingreso familiar; NEP: nivel educativo del padre; NEM: nivel educativo de la madre.

La Tabla II muestra que existe una correlación positiva, moderada y significativa entre FC y ECS, ECA, ECE y AA. La variable con mayor correlación con FC fue ECS; la variable con menor correlación es AA. Además, existe una correlación positiva, de nivel moderado y significativa entre PRCM y IF, NEP y NEM, y una correlación negativa, de nivel bajo y significativa con G. La variable con mayor correlación con PRCM es NEM; la variable con menor correlación es G.

### Resultados para predictores *continuas*

El orden de importancia y los coeficientes estandarizados obtenidos con RM y RNA en condiciones donde hay 2, 3 y 4 predictores *continuas* en el conjunto de datos real (conjunto de datos1) se presentan en la Tabla III.

**TABLA III.** Orden de importancia de RM y RNA en el conjunto de datos real que consta de predictores *continuas*

VD	Método	NP	Orden de importancia de los predictores y coeficientes estandarizados (entre paréntesis)							
FC	RM	2	1.	ECS (0.34)	2.	ECA (0.32)				
		3	1.	ECA (0.27)	2.	ECS (0.26)	3.	ECE (0.22)		
		4	1.	ECS (0.25)	2.	ECA (0.23)	3.	ECE (0.16)	4.	AA (0.16)
	RNA	2	1.	ECS (0.50)	2.	ECA (0.49)				
		3	1.	ECA (0.35)	2.	ECS (0.34)	3.	ECE (0.31)		
		4	1.	ECS (0.30)	2.	ECA (0.29)	3.	ECE (0.21)	4.	AA (0.20)

Nota. VD: variable dependiente; NP: número de predictores.

La Tabla III muestra que para las condiciones con dos y cuatro predictores en la predicción de FC, ECS es la variable con mayor nivel de importancia según ambos métodos. En la predicción con tres variables, aunque el orden de importancia de las variables predictoras es el mismo en cuanto a métodos, la variable con mayor nivel de importancia es ECA. Estos resultados muestran que ambos métodos revelan el mismo orden para la condición donde las variables predictoras son *continuas*, independientemente del número de predictores. El orden de importancia y los coeficientes estandarizados obtenidos por RM a partir de conjuntos de datos de simulación con el diferente número de predictores y tamaños de muestra se presentan en la Tabla IV.

**TABLA IV.** Orden de importancia de RM y coeficientes estandarizados en conjuntos de datos de simulación que consisten en predictores *continuas*

VD	NP	TM	Orden de importancia de los predictores y coeficientes estandarizados (entre paréntesis)							
FC	2	100	1.	ECS (0.34)	2.	ECA (0.32)				
		200	1.	ECS (0.34)	2.	ECA (0.31)				
		400	1.	ECS (0.34)	2.	ECA (0.32)				
		800	1.	ECS (0.34)	2.	ECA (0.32)				
	3	100	1.	ECA (0.27)	2.	ECS (0.26)	3.	ECE (0.23)		
		200	1.	ECS (0.27)	2.	ECA (0.26)	3.	ECE (0.22)		
		400	1.	ECA (0.27)	2.	ECS (0.25)	3.	ECE (0.23)		
		800	1.	ECA (0.27)	2.	ECS (0.26)	3.	ECE (0.23)		
	4	100	1.	ECS (0.24)	2.	ECA (0.23)	3.	ECE (0.18)	4.	AA (0.17)
		200	1.	ECS (0.26)	2.	ECA (0.22)	3.	ECE (0.17)	4.	AA (0.16)
		400	1.	ECS (0.25)	2.	ECA (0.23)	3.	ECE (0.17)	4.	AA (0.16)
		800	1.	ECS (0.25)	2.	ECA (0.23)	3.	ECE (0.17)	4.	AA (0.16)

Nota. TM: tamaño de la muestra.

La tabla IV muestra que la variable con mayor nivel de significación según los resultados de RM obtenidos de dos predictores en diferentes tamaños de muestra es ECS. En la condición de que el número de predictores sea cuatro, la variable con mayor nivel de importancia vuelve a ser ECS. Además, en condiciones con dos y cuatro predictores, el orden de importancia de los predictores es el mismo para todos los tamaños de muestra. Por otro lado, en las condiciones con tres predictores, ECS

es la variable con mayor nivel de importancia para el tamaño de muestra de 200; es ECA para tamaños de muestra de 100, 400 y 800. Además, para esta condición, ECE es la variable menos importante en todos los tamaños de muestra. El orden de importancia obtenido para todos los predictores se ajusta generalmente a los datos reales (Tabla III). El orden de importancia y los coeficientes estandarizados obtenidos por RNA a partir de conjuntos de datos de simulación con diferente número de predictores y tamaños de muestra se presentan en la Tabla V.

**TABLA V.** El orden de importancia de RNA y coeficientes estandarizados en conjuntos de datos de simulación que consisten en predictores *continuas*

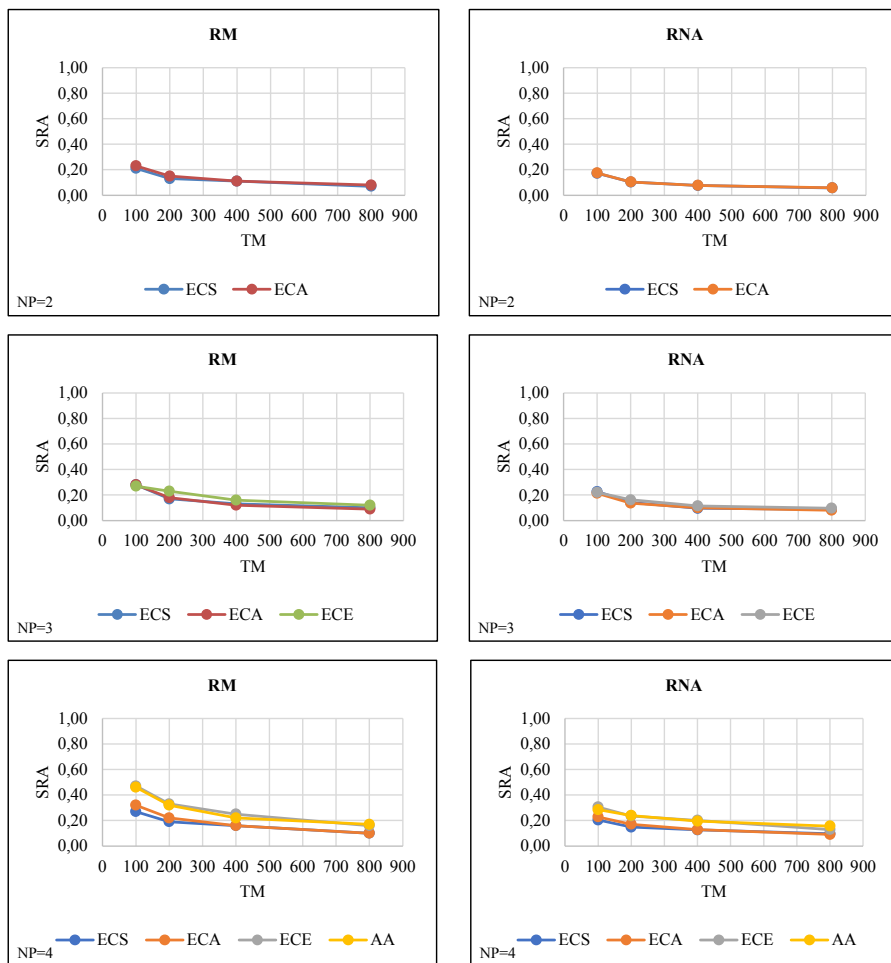
VD	NP	TM	Orden de importancia de los predictores y coeficientes estandarizados (entre paréntesis)							
FC	2	100	1.	ECS (0.52)	2.	ECA (0.48)				
		200	1.	ECS (0.52)	2.	ECA (0.48)				
		400	1.	ECS (0.52)	2.	ECA (0.48)				
		800	1.	ECS (0.51)	2.	ECA (0.49)				
	3	100	1.	ECS (0.35)	2.	ECA (0.34)	3.	ECE (0.31)		
		200	1.	ECS (0.35)	2.	ECA (0.34)	3.	ECE (0.31)		
		400	1.	ECA (0.35)	2.	ECS (0.34)	3.	ECE (0.31)		
		800	1.	ECA (0.35)	2.	ECS (0.34)	3.	ECE (0.31)		
	4	100	1.	ECS (0.27)	2.	ECA (0.26)	3.	ECE (0.24)	4.	AA (0.22)
		200	1.	ECS (0.30)	2.	ECA (0.26)	3.	ECE (0.23)	4.	AA (0.21)
		400	1.	ECS (0.29)	2.	ECA (0.27)	3.	ECE (0.22)	4.	AA (0.21)
		800	1.	ECS (0.30)	2.	ECA (0.28)	3.	ECE (0.21)	4.	AA (0.20)

La Tabla V muestra que ECS es la variable con mayor nivel de importancia según los resultados de RNA obtenidos de dos y cuatro predictores en diferentes tamaños de muestra. Además, en condiciones con dos y cuatro predictores, el orden de importancia de los predictores es el mismo para todos los tamaños de muestra. Por otro lado, en la condición con tres predictores, ECS es la variable con mayor nivel de importancia para tamaños de muestra de 100 y 200; para tamaños de muestra de 400 y 800 es ECA. Además, para esta condición, ECE es la variable menos importante en todos los tamaños de muestra. El orden de importancia obtenido para todos los predictores se ajusta generalmente a los datos reales (Tabla III).

Los SRAs de las estimaciones de parámetros obtenidas por RM y RNA a partir de conjuntos de datos de simulación con diferentes números de predictores *continuas* y tamaños de muestra se presentan en la Figura I.



FIGURA I. SRAs obtenidos por RM y RNA en conjuntos de datos de simulación que consisten en predictores *continuas*



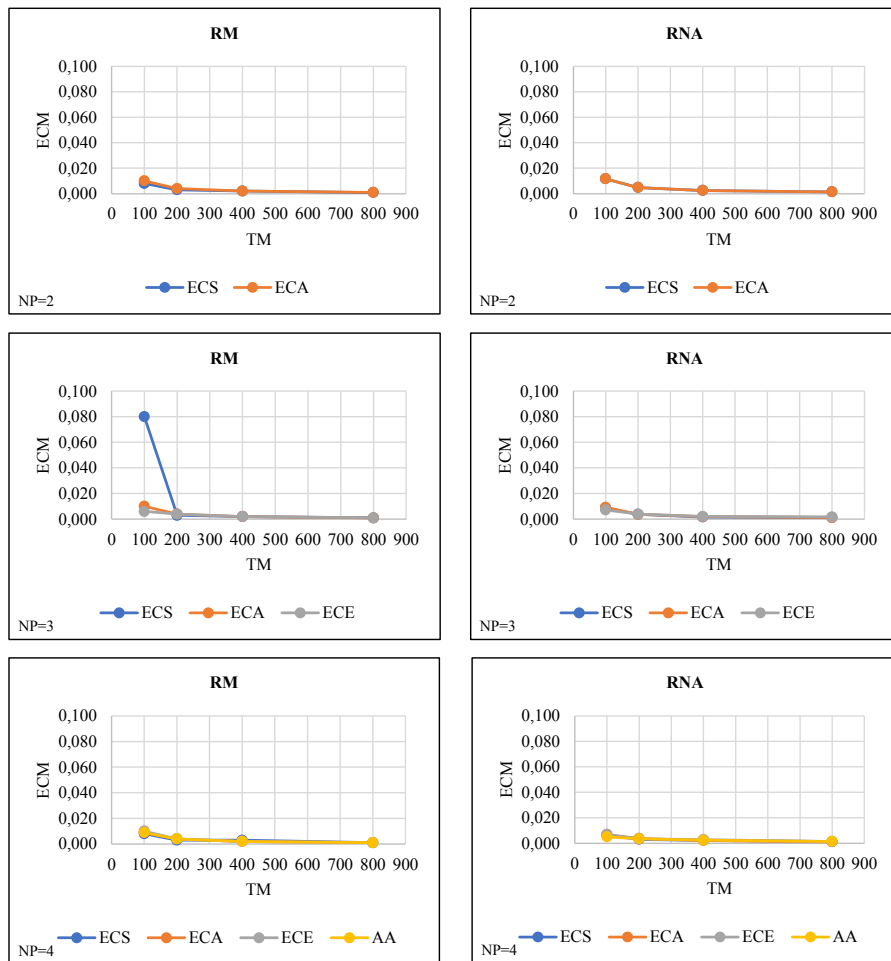
La Figura I muestra que SRA disminuye a medida que aumenta el tamaño de la muestra para ambos métodos, independientemente del número de predictores. Si bien se observaron disminuciones dramáticas en SRA, específicamente en la transición de 100 a 200 tamaños de muestra, estas disminuciones fueron menores en muestras más grandes

(es decir, 400-800). Además, para todos los predictores, los SRAs más altos se observaron con un tamaño de muestra de 100 y los SRAs más bajos con un tamaño de muestra de 800. Cuando se compararon las medias de RM y RNA SRA, se determinó que RNA tenía promedios SRA más bajos, independientemente del número de predictores. Sin embargo, dado que los promedios SRA de ambos métodos superan el 10% para todos los números de predictores (Flora & Curran, 2004), se puede decir que estas estimaciones tienen un sesgo sustancial.

Por otro lado, el aumento en el número de predictores mientras se mantuvo fijo el tamaño de la muestra, provocó un aumento en las medias SRA de ambos métodos. Sin embargo, el aumento en las medias de RNA SRA fue menor en comparación con RM. Cabe señalar que estos aumentos en las medias de SRA fueron más dramáticos, específicamente en tamaños de muestra pequeños.

Los ECMs de las estimaciones de parámetros obtenidos por RM y RNA a partir de conjuntos de datos de simulación con diferentes números de predictores *continuas* y los tamaños de muestra se presentan en la Figura II.

FIGURA II. ECMs obtenidos por RM y RNA en conjuntos de datos de simulación que consisten en predictores *continuos*



La figura II muestra que para ambos métodos, independientemente del número de predictores, el ECM disminuye a medida que aumenta el tamaño de la muestra. En general, se observaron disminuciones dramáticas en ECM en transiciones de 100 a 200 tamaños de muestra; en muestras más grandes (es decir, 400-800), estas disminuciones fueron

menores. Además, para todos los predictores, los valores de ECM más altos se observaron con un tamaño de muestra de 100 y los ECMs más bajos con un tamaño de muestra de 800. Cuando se compararon los ECMs de RM y RNA, se determinó que RNA tenía ECMs más bajos, excepto cuando el número de predictores era dos.

Por otro lado, a medida que aumentó el número de predictores, no se observó un cambio importante en los ECMs de los métodos (excepto NP=3, ECS). Además, se puede decir que los ECMs tienen valores bastante similares en muestras de 200 y más.

### Resultados para predictores discretos

El orden de importancia y los coeficientes estandarizados obtenidos con RM y RNA en condiciones donde hay 2, 3 y 4 predictores *discretos* en el conjunto de datos reales se presentan en la Tabla VI.

**TABLA VI.** Orden de importancia de RM y RNA en el conjunto de datos real que consta de predictores *discretos*

VD	Método	NP	Orden de importancia de los predictores y coeficientes estandarizados (entre paréntesis)							
PRCM	RM	2	1.	IF (0.25)	2.	G (0.21)				
		3	1.	G (0.19)	2.	IF (0.17)	3.	NEP (0.13)		
		4	1.	G (0.19)	2.	IF (0.14)	3.	NEP (0.09)	4.	NEM (0.09)
	RNA	2	1.	IF (0.27)	2.	G (0.20)				
		3	1.	G (0.16)	2.	IF (0.15)	3.	NEP (0.13)		
		4	1.	G (0.14)	2.	IF (0.10)	3.	NEM (0.09)	4.	NEP (0.08)

La Tabla VI muestra que para la condición con dos predictores en la predicción de PRCM, el predictor con mayor nivel de importancia según ambos métodos fue IF y que G para condiciones con tres y cuatro predictores. Además, para las condiciones donde el número de predictores es 2 y 3, el orden de importancia obtenido por ambos métodos es el mismo; se vio que solo el tercer y cuarto orden cambiaron

cuando el número de predictores fue 4. El orden de importancia y los coeficientes estandarizados obtenidos por RM a partir de conjuntos de datos de simulación con el diferente número de predictores y tamaños de muestra se presentan en la Tabla VII.

**TABLA VII.** Orden de importancia de RM y coeficientes estandarizados en conjuntos de datos de simulación que consisten en predictores *discretos*

VD	NP	TM	Orden de importancia de los predictores y coeficientes estandarizados (entre paréntesis)							
PRCM	2	100	1.	IF (0.28)	2.	G (0.18)				
		200	1.	IF (0.28)	2.	G (0.21)				
		400	1.	IF (0.26)	2.	G (0.20)				
		800	1.	IF (0.27)	2.	G (0.19)				
	3	100	1.	IF (0.21)	2.	NEP (0.20)	3.	G (0.17)		
		200	1.	IF (0.20)	2.	G (0.19)	3.	NEP (0.19)		
		400	1.	G (0.19)	2.	NEP (0.18)	3.	IF (0.18)		
		800	1.	NEP (0.19)	2.	IF (0.19)	3.	G (0.18)		
	4	100	1.	IF (0.18)	2.	NEP (0.17)	3.	G (0.17)	4.	NEM (0.15)
		200	1.	G (0.19)	2.	IF (0.17)	3.	NEP (0.14)	4.	NEM (0.12)
		400	1.	G (0.19)	2.	IF (0.15)	3.	NEP (0.13)	4.	NEM (0.12)
		800	1.	G (0.18)	2.	IF (0.15)	3.	NEP (0.13)	4.	NEM (0.12)

La Tabla VII muestra que IF es el predictor con mayor importancia de acuerdo con los resultados de RM obtenidos de dos predictores en diferentes tamaños de muestra. Cuando el número de predictores era tres, se diferenciaba el predictor con mayor nivel de importancia según los tamaños de muestra. En el caso de que el número de predictores sea cuatro, el predictor con el nivel de importancia más alto en tamaños de muestra de 200 y superiores es G. Mientras que las clasificaciones de

importancia obtenidas de las condiciones con dos y cuatro predictores fueron generalmente congruentes con las clasificaciones de importancia de los datos reales (Tabla VI); no se observó tal concordancia en condiciones en las que el número de predictores era tres. El orden de importancia y los coeficientes estandarizados obtenidos por RNA a partir de conjuntos de datos de simulación con diferentes números de predictores y tamaños de muestra se presentan en la Tabla VIII.

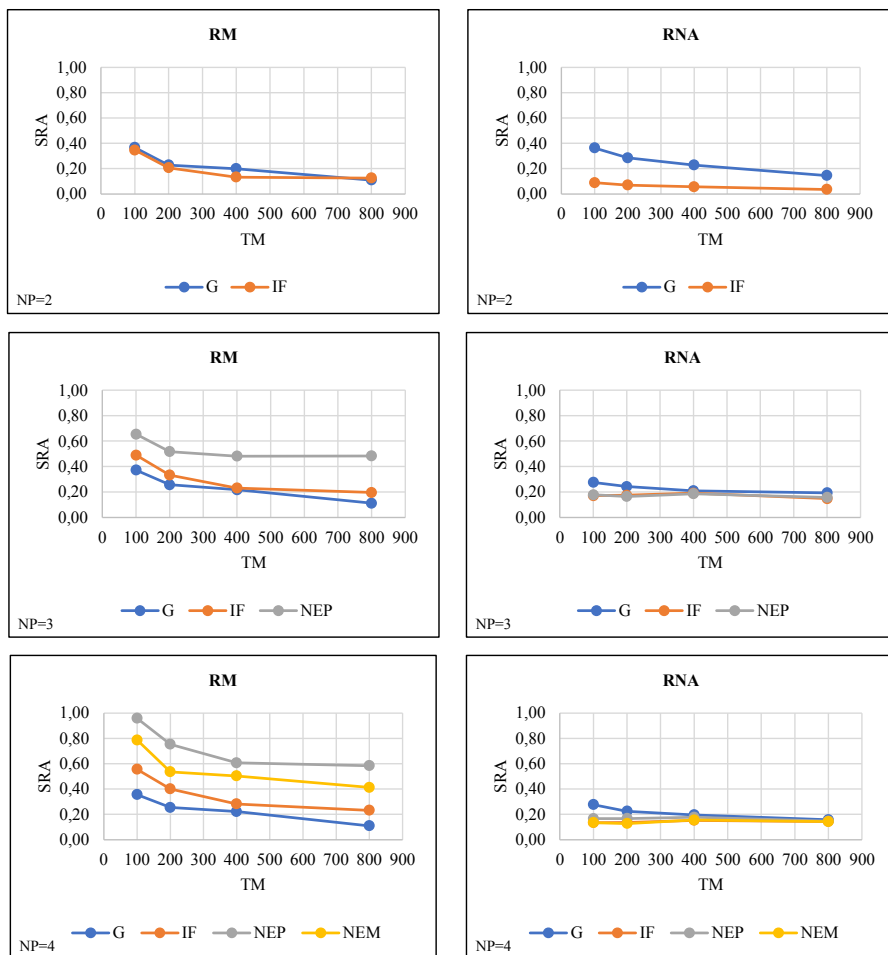
**TABLA VIII.** Orden de importancia de RNA y coeficientes estandarizados en conjuntos de datos de simulación que consisten en predictores *discretos*

VD	NP	TM	Orden de importancia de los predictores y coeficientes estandarizados (entre paréntesis)							
PRCM	2	100	1.	IF (0.26)	2.	G (0.21)				
		200	1.	IF (0.26)	2.	G (0.22)				
		400	1.	IF (0.26)	2.	G (0.21)				
		800	1.	IF (0.27)	2.	G (0.19)				
	3	100	1.	IF (0.16)	2.	G (0.14)	3.	NEP (0.14)		
		200	1.	G (0.16)	2.	IF (0.15)	3.	NEP (0.14)		
		400	1.	G (0.15)	2.	IF (0.15)	3.	NEP (0.14)		
		800	1.	IF (0.16)	2.	NEP (0.14)	3.	G (0.13)		
	4	100	1.	G (0.11)	2.	IF (0.10)	3.	NEM (0.09)	4.	NEP (0.09)
		200	1.	G (0.13)	2.	IF (0.10)	3.	NEP (0.09)	4.	NEM (0.09)
		400	1.	G (0.13)	2.	IF (0.10)	3.	NEM (0.09)	4.	NEP (0.09)
		800	1.	G (0.12)	2.	IF (0.10)	3.	NEM (0.09)	4.	NEP (0.09)

La Tabla VIII muestra que IF es el predictor con mayor nivel de importancia según los resultados de RNA obtenidos de diferentes tamaños de muestra con dos predictores. Cuando el número de predictores era tres, se diferenciaba el predictor con mayor nivel de importancia según los tamaños de muestra. Para la condición en la que el número de predictores era cuatro, el predictor con el nivel de importancia más alto es G. Cuando se tienen en cuenta todas las condiciones (excepto NP=3, TM=100 y 800; NP=4, TM=200), las clasificaciones de importancia predictiva obtenidas de los conjuntos de datos de simulación fueron generalmente congruentes con las clasificaciones de importancia de los datos reales (Tabla VI).

Los SRAs de las estimaciones de parámetros obtenidos por RM y RNA a partir de conjuntos de datos de simulación con diferentes números de predictores *discretos* y los tamaños de muestra se presentan en la Figura III.

FIGURA III. SRAs obtenidos por RM y RNA en conjuntos de datos de simulación que consisten en predictores *discretos*



La Figura III muestra que el SRA generalmente disminuye a medida que aumenta el tamaño de la muestra para ambos métodos, independientemente del número de predictores. Además, los SRAs RNA se vieron menos afectados por el aumento del tamaño de la muestra en comparación con RM. En general, se observaron disminuciones dramáticas en SRA, específicamente en la transición de 100 a 200

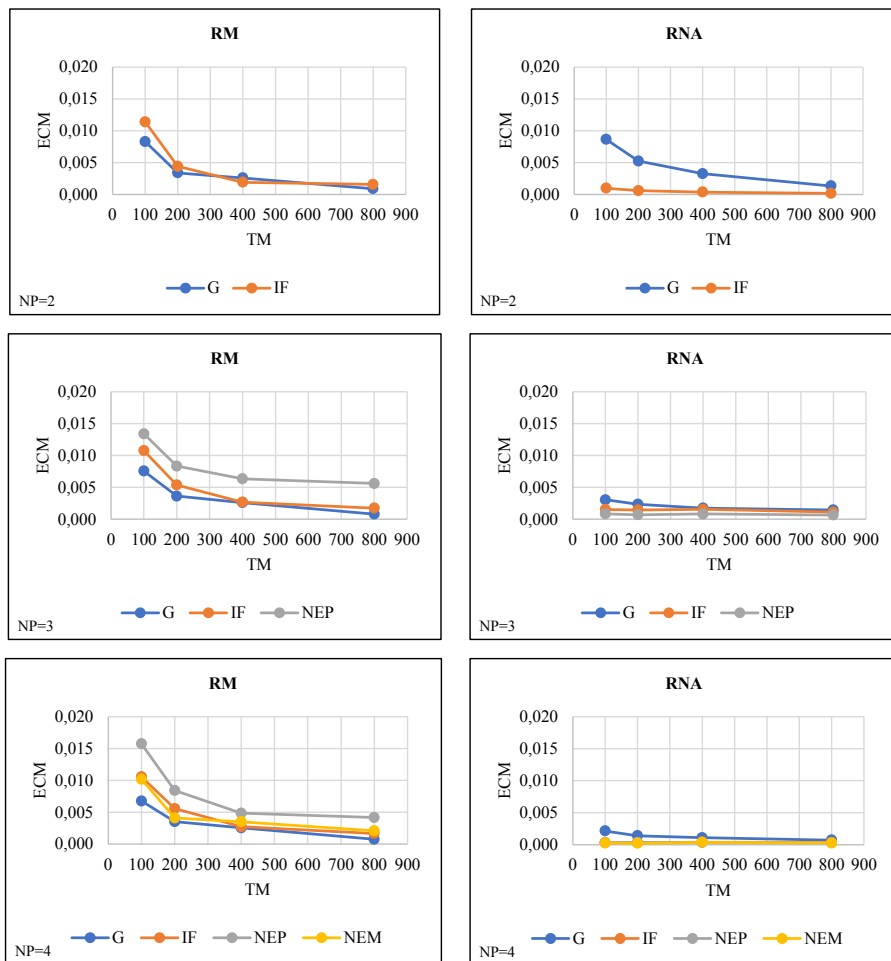


tamaños de muestra, estas disminuciones fueron menores en muestras más grandes (es decir, 400-800). Además, para todos los predictores en RM, los SRAs más altos se observaron con un tamaño de muestra de 100 y los SRAs más bajos con un tamaño de muestra de 800. Por otro lado, no se detectó un patrón común relacionado con el aumento de los tamaños de muestra en los SRAs RNA. Por ejemplo, mientras que el aumento en el tamaño de la muestra corrige las SRAs del predictor de género, no tuvo el mismo efecto en las SRAs de otras variables (IF, NEP, NEM). Cuando se compararon las medias de RM y RNA SRA, se determinó que RNA tenía SRAs más bajas, independientemente del número de predictores. Dado que los promedios SRA de ambos métodos superan el 10% para todos los números predictores (Flora & Curran, 2004), se puede decir que estas estimaciones tienen un sesgo sustancial.

Por otro lado, cuando el tamaño de la muestra es fijo, el aumento en el número de predictores condujo a un aumento en las medias de RM SRA; no se observó ningún patrón común (es decir, aumento, disminución) para las medias RNA SRA. Además, llama la atención que el deterioro de las medias de SRA es más dramático, específicamente en pequeños tamaños de muestra.

Los ECMs de las estimaciones de parámetros obtenidos por RM y RNA a partir de conjuntos de datos de simulación con diferentes números de predictores *discretos* y los tamaños de muestra se presentan en la Figura IV.

FIGURA IV. ECMs obtenidos por RM y RNA en conjuntos de datos de simulación que consisten en predictores *discretos*



La Figura IV muestra que el ECM generalmente disminuye a medida que aumenta el tamaño de la muestra para ambos métodos, independientemente del número de predictores. Además, las RNA ECMs se vieron menos afectadas por el aumento del tamaño de la muestra en comparación con RM. En general, se observaron disminuciones dramáticas en los ECMs del RM, específicamente en la transición de 100 a 200 tamaños de muestra, estas disminuciones fueron menores en

muestras más grandes (es decir, 400-800). Por otro lado, las ECMs de RNA generalmente no se vieron afectadas por el aumento en el tamaño de la muestra. Además, se puede decir que para todos los predictores en ambos métodos, los ECMs más altos se observaron en un tamaño de muestra de 100, y los ECMs más bajos se observaron en un tamaño de muestra de 800. Cuando se compararon los ECMs de RM y RNA, se determinó que RNA generalmente tenía ECMs más bajos, independientemente del número de predictores.

En general, RNA ECMs disminuyó mientras que RM ECMs aumentó a medida que aumentaba el número de predictores. Además, cabe señalar que el deterioro de las ECMs RM es más dramático, específicamente en tamaños de muestra pequeños.

## Discusión

En este estudio, se compararon los rendimientos de RM y RNA para determinar el orden de importancia de las variables predictoras. Para ello, se utilizaron conjuntos de datos reales y datos de simulación. El tipo de variable, el tamaño de la muestra y el número de predictores fueron manipulados en el estudio de simulación. Primero, se examinaron las correlaciones entre las variables en los conjuntos de datos reales y los supuestos de RM. Luego, se compararon los desempeños de los métodos sobre el orden de importancia de sus predictores, SRA y ECM.

## Comparación de clasificación de importancia de los métodos predictores

Cuando las clasificaciones de importancia de los predictores se obtienen del conjunto de datos real en el que se comparan los predictores *continuas*, se ve que las clasificaciones de ambos métodos son las mismas independientemente del número de predictores. Las clasificaciones de importancia obtenidas de los conjuntos de datos de simulación son las mismas que las clasificaciones de importancia de los datos reales, excepto por una condición de simulación para RM y dos para RNA.

Por otro lado, cuando se comparan las clasificaciones de importancia de los predictores obtenidos del conjunto de datos real en el que los predictores son *discretos* las clasificaciones de importancia obtenidos por

ambos métodos son los mismos para las condiciones donde el número de predictores es dos y tres. Con la condición de que el número de predictores sea cuatro, solo han cambiado los rangos tercero y cuarto. Dado que este cambio es causado por un valor inferior a 0.01, se puede afirmar que no es muy importante. Por lo tanto, se puede decir que los métodos tienen casi las mismas clasificaciones. Las clasificaciones de importancia obtenidas de los conjuntos de datos de simulación son las mismas que las clasificaciones de importancia de los datos reales, excepto por cinco condiciones de simulación para RM y tres para RNA. Además, cabe señalar que es posible que RM no realice clasificaciones de importancia de datos reales cuando el número de predictores es tres.

En resumen, en el presente estudio se puede decir que los métodos desempeñan el mismo orden de importancia si los predictores son *continuas* y desempeñan un orden de importancia bastante similar si los predictores son *discretos*. Turhan et al. (2013) realizaron RM y RNA al conjunto de datos que constaba de predictores *continuas* y *discretos* e informaron que la misma variable fue el mejor predictor según ambos métodos. Se puede afirmar que este hallazgo apoya los hallazgos de la presente investigación.

### Comparación de SRAs de métodos

En condiciones en las que los predictores son *continuas* y *discretos*, a medida que aumenta el tamaño de la muestra, el SRA de los métodos disminuye independientemente del número de predictores. Este efecto correctivo del aumento en el tamaño de la muestra es específicamente mayor cuando se pasa de 100 a 200. Además, cuando se comparan las medias de SRA, se puede decir que RNA supera a RM, independientemente del número de predictores. Por otro lado, cabe señalar que los métodos realizan predicciones con un sesgo sustancial (>10%; Flora & Curran, 2004).

Por otro lado, en condiciones donde los predictores son *continuas*, mientras que el tamaño de la muestra es fijo, el aumento en el número de predictores conduce a un aumento en las medias SRA de los métodos. En general, el aumento en las medias de RM SRA es mayor en comparación con RNA. En condiciones en las que los predictores son *discretos*, mientras que el tamaño de la muestra es fijo, la media del RM SRA aumenta a

medida que aumenta el número de predictores. Sin embargo, tal patrón común (es decir, aumento, disminución) no se determina en RNA. Además, los aumentos en la media de SRA son mayores, específicamente en tamaños de muestra pequeños, en condiciones donde los predictores son tanto *continuas* como *discretos*. Teniendo en cuenta los medios SRA, se puede afirmar que RNA supera a RM.

## Comparación de ECMs de métodos

En condiciones en las que los predictores son tanto *continuas* como *discretos*, los ECMs de los métodos tienden a disminuir a medida que aumenta el tamaño de la muestra, independientemente del número de predictores. Este efecto remediador del aumento del tamaño de la muestra alcanza su nivel máximo, específicamente cuando se aumenta el tamaño de la muestra de 100 a 200. Además de esto, es bastante bajo en otras transiciones de muestra. Cuando se comparan los ECMs medios de los métodos, se puede decir que RNA generalmente supera a RM. Lykourentzou et al. (2009) y Turhan et al. (2013), quienes compararon diferentes variables que predicen el rendimiento de los estudiantes en el campo de las ciencias de la educación, revelaron que RNA realiza mejores predicciones que RM. Altún et al. (2019) estimaron las calificaciones de graduación de los estudiantes de enseñanza en el aula utilizando RM y RNA y determinaron que ambos métodos (RM=94.30%, RNA=94.43%) proporcionaron resultados muy parecidos. Por otro lado, cuando se examinan los estudios que comparan los métodos en el campo de los negocios sobre las ECMs, Akbilgiç y Kesintürk (2008); Aktaş et al. (2003) y Yüzük (2019) informaron que RNA superó a RM. Cuando se examinan los estudios que comparan los métodos en el campo de la ingeniería sobre las ECMs, Okkan y Mollamahmutoğlu (2010) afirman que RNA supera a RM, mientras que Cansız et al. (2020) informaron que RM superó a RNA.

Por otro lado, cuando el número de predictores aumenta mientras el tamaño de la muestra es fijo, no se determina un cambio importante en los ECMs de los métodos en las condiciones donde los predictores son *continuas*. Los ECMs de los métodos son bastante similares, específicamente en el tamaño de la muestra de 200 o más. En las condiciones en las que los predictores son *discretos*, cuando el número

de predictores aumenta mientras el tamaño de la muestra es fijo, los RM ECMs aumentan en general; RNA ECMs, por otro lado, tienden a disminuir. Cabe señalar que los aumentos en las ECMs de RM son más drásticos, específicamente en tamaños de muestra pequeños.

## Conclusión y Sugerencias

Generalmente, el aumento en el tamaño de la muestra, independientemente del tipo de predictor (es decir, *continua* y *discreto*) y el número de predictores, remedia los SRAs y ECMs de los métodos. Para obtener estimaciones óptimas con ambos métodos, se recomienda utilizar un tamaño de muestra de al menos 200. Sin embargo, cabe señalar que las estimaciones obtenidas para este tamaño de muestra pueden tener un sesgo moderado.

En caso de que los predictores sean *continuas*, los investigadores pueden preferir uno de los dos métodos ya que el orden de importancia del predictor de los métodos es el mismo. Sin embargo, en los casos en que los predictores sean *discretos* y el número de predictores aumente (específicamente >2), se puede recomendar RNA debido a su rendimiento superior.

Por otro lado, cuando los predictores son *continuas*, el aumento en el número de predictores conduce a un deterioro de las medias SRA de los métodos. Sin embargo, cabe señalar que este deterioro es más dramático en la RM. En caso de que los predictores sean *discretos*, RNA proporciona estimaciones más sólidas que RM.

Si bien los predictores son *continuas*, el aumento en el número de predictores no conduce a un cambio importante en los ECMs de los métodos. Los ECMs de los métodos son bastante similares, específicamente en el tamaño de la muestra de 200 o más. Cuando los estimadores son *discretos*, el aumento en el número de estimadores generalmente conduce a un deterioro en los ECMs de la RM, mientras que los ECMs de la RNA se ven afectados levemente.

Como resultado, cuando se consideran SRA y ECMs, RNA supera a RM. Por esta razón, se puede sugerir que los investigadores prefieran RNA en primer lugar. RM es una alternativa razonable a RNA si los predictores son *continuas* y se cumplen sus suposiciones. Sin embargo, debe tenerse en cuenta que es posible que RM no funcione bien incluso si se cumplen

sus suposiciones si los predictores son *discretos*. Específicamente, en el caso de tres o más predictores *discretos* se puede recomendar que los investigadores utilicen RNA. Además, se debe enfatizar que una ventaja importante de RNA es que no necesita los supuestos requeridos por RM.

Es importante que la investigación educativa revele métodos de estimación que permitan determinar el orden de importancia de los predictores en diferentes modelos. Por lo tanto, se puede suponer que los estudios en los que se abordan comparativamente los métodos tienen el potencial de contribuir a otros estudios en el campo de la educación. En este contexto, los resultados del presente estudio brindan una mejor comprensión del funcionamiento de RM y RNA, lo que contribuye a la investigación educativa, en diferentes condiciones.

La presente investigación se ha llevado a cabo sobre dos conjuntos de datos reales diferentes donde se cumplen los supuestos de la RM y las variables predictoras son *continuas* o *discretos*. Los datos de simulación se generaron considerando las relaciones en estos conjuntos de datos. Este diseño de investigación puede considerarse como una limitación de la presente investigación. Además, cabe señalar que otra limitación es que los predictores explican la diferencia total de la variable dependiente en los conjuntos reales, que es del 34.3% y 20.4%, respectivamente. Hasta donde sabemos, dado que no hay estudios que comparen ambos métodos en el campo de la educación bajo condiciones como el tipo de variable, el tamaño de la muestra y el número de predictores, se necesita más investigación para generalizar los resultados. En estudios futuros, los datos se pueden generar modelando relaciones en muestras más grandes, y las comparaciones de métodos se pueden realizar a través de modelos en los que se usan predictores combinados *continuas* y *discretos*.

## Referencias bibliográficas

Akbilgiç, O., & Keskindürk, T. (2008). The comparison of artificial neural networks and regression analysis. *Istanbul Management Journal*, 19(60), 74-83.

- Akgül, S. (2019). Predictive power of mathematical self-efficacy for gifted and talented students' mathematical achievement. *EKEV Journal*, 23(78), 481-496.
- Akhtar, M., & Herwig, B. K. (2019). Coping styles and socio-demographic variables as predictors of psychological well-being among international students belonging to different cultures. *Curr Psychol*, 38, 618-626.
- Aktaş, R., Doğanay, M., & Yıldız, B. (2003). Financial failure prediction: Statistical methods and artificial neural network comparison. *Ankara University SBF Journal*, 58, 1-24.
- Alandete, J. G. (2015). Does meaning in life predict psychological well-being? An analysis using the Spanish versions of the purpose-in-life test and the Ryff's scales. *The European Journal of Counselling Psychology*, 3(2), 89-98.
- Amatya, A., & Demirtas, H. (2015). OrdNor: An R package for concurrent generation of correlated ordinal and normal data. *Journal of Statistical Software*, 68, 1-14.
- Altun, M., Kayıkçı, K., & Irmak, S. (2019). Estimation of graduation grades of primary education students by using regression analysis and artificial neural networks. *e-IJER*, 10(3), 29-43.
- Aramburo, V., Boroel, B., & Pineda, G. (2017). Predictive factors associated with academic performance in college students. *Social and Behavioral Sciences*, 237, 945-949.
- Arthur, W., & Doverspike, D. (2001). Predicting motor vehicle crash involvement from a personality measure and a driving knowledge test. *Journal of Prevention & Intervention in the Community*, 22(1), 35-42.
- Atabey, N. (2020). Future expectations and self-efficacy of high school students as a predictor of sense of school belonging. *Education and Science*, 45(201), 125-141.
- Bandalos, D. L., & Leite, W. (2013). The use of Monte Carlo studies in structural equation modeling research. In G. R. Hancock & R. O. Mueller (Eds.), *Structural equation modeling: A second course* (pp. 385-426). Information Age Publishing.
- Baş, N. (2006). *Artificial neural networks approach and an application* (Unpublished master's thesis). Mimar Sinan Fine Arts University.
- Bayru, P. (2007). *Electronic media consumer choice analysis: Artificial neural networks to evaluate the performance of the model with lojit* (Unpublished doctoral dissertation). Istanbul University.



- Bedel, A., Işık, E., & Hamarta, E. (2014). Psychometric properties of the KIDCOPE in Turkish adolescents. *Education and Science*, 39(176), 227-235.
- Bergold, S., & Steinmayr, R. (2018). Personality and intelligence interact in the prediction of academic achievement. *Journal of Intelligence*, 6(2), 27.
- Büyüköztürk, Ş. (2014). *Sosyal bilimler için veri analizi el kitabı [Manual of data analysis for social sciences]*. Pegem Yayıncılık.
- Cansız, Ö. F., Öztekin, N., & Erginer, İ. (2020). Comparison of artificial neural networks and multi-variable linear regression techniques in determination of the number of optimum vehicles. *DUMF Journal*, 11(2), 771-782.
- Chalmers, R. P., & Adkins, M. C. (2020). Writing effective and reliable Monte Carlo simulations with the SimDesign package. *The Quantitative Methods for Psychology*, 16(4), 248-280.
- Çelikkaleli, Ö. (2014). The validity and reliability of the cognitive flexibility scale. *Education and Science*, 39(176), 339-346.
- Çelikkaleli, Ö., Gündoğdu, M., & Kıran-Esen, B. (2006). Questionnaire for measuring self-efficacy in youths: Validity and reliability study of Turkish form. *EJER*, 25, 62-72.
- Çırak, G., & Çokluk, Ö. (2013). The usage of artificial neural network and logistic regression methods in the classification of student achievement in higher education. *MJH*, 3(2), 71-79.
- Elmas, Ç. (2003). *Yapay sinir ağları [Artificial neural networks]*. Seçkin Kitabevi.
- Flora, D. B., & Curran, P. J. (2004). An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. *Psychological Methods*, 9(4), 466-491.
- Fraenkel, J. R., Wallen, N., & Hyun, H. (2011). *How to design and evaluate research in education*. McGraw Hill.
- Garcini, L. M., Short, M. B., & Norwood, W. D. (2013). Affective and motivational predictors of perceived meaning in life among college students. *The Journal of Happiness and Well-Being*, 1(2), 47-60.
- Gorr, W. L., Nagin, D., & Szczypula, A. (1994). Comparative study of artificial neural network and statistical models for predicting student grade point averages. *International Journal of Forecasting*, 10(2), 17-34.

- Haykin, S. (1994). *Neural networks: A comprehensive foundation*. Macmillan Press.
- IBM. (2020). IBM SPSS statistics base 25. SPSS Inc.
- Ivcevic, Z., & Eggers, C. (2021). Emotion regulation ability: Test performance and observer reports in predicting relationship, achievement and well-being outcomes in adolescents. *IJERPH*, 18(6), 3204.
- Kalkan, Ö. K., & Coşguner, T. (2021). Evaluation of the academic achievement of vocational school of higher education students through artificial neural networks. *Gazi University Journal of Science*, 34(3), 851-862.
- Lee, J. (2016). Attitude toward school does not predict academic achievement. *Learning and Individual Differences*, 52, 1-9.
- Lykourantzou, I., Giannoukos, I., Mpardis, G., Nikolopoulos, V., & Loumos, V. (2009). Early and dynamic student achievement prediction in e-learning courses using neural networks. *JASIST*, 60(2), 372-380.
- Marroquin, B., & Nolen-Hoeksema, S. (2015). Event prediction and affective forecasting in depressive cognition: Using emotion as information about the future. *Journal of Social and Clinical Psychology*, 34(2), 117-134.
- Martin, M. M., & Rubin, R. B. (1995). A new measure of cognitive flexibility. *Psychological Reports*, 76, 623-626.
- Moreetsi, T., & Mbako, M. T. (2008). Predicting students' performance on agricultural science examination from forecast grades. *US-China Education Review*, 5(10), 45-51.
- Muris, P. (2001). A brief questionnaire for measuring self-efficacy in youths. *Journal of Psychopathology and Behavioral Assessment*, 23(3), 145-149.
- Okkan, U., & Mollamahmutoğlu, A. (2010). Daily runoff modeling of Yiğitler stream by using artificial neural networks and regression analysis. *Dumlupınar University Journal of Science Institute*, 23, 33-48.
- Öztemel, E. (2012). *Yapay sinir ağları [Artificial neural networks]*. Papatya Yayıncılık.
- Pedhazur, E. J. (1997). *Multiple Regression in behavioral research, explanation and prediction*. Thomson Learning.
- R Core Team. (2017). *R: A language and environment for statistical computing [Computer Software]*. R Foundation for Statistical Computing.

- Simpson, P. K. (1990). *Artificial neural systems foundations, paradigms, application and implementation*. Pergamon Press.
- Spirito, A., Stark, L. J., & Williams, C. (1988). Development of a brief coping checklist for use with pediatric populations. *Journal of Pediatric Psychology*, 13, 555-574.
- Stevens, N. (2009). *Applied multivariate statistics for the social sciences*. Taylor and Francis.
- Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics*. Pearson.
- Teressa, T. D., & Bekele, G. (2021). Motivational predictors of tenth graders' academic achievement in Harari secondary schools. *IJELS*, 9(3), 9-19.
- Tolon, M. (2007). *Measuring customer satisfaction with artificial neural networks and an application of retail consumers in Ankara* (Unpublished doctoral dissertation). Gazi University.
- Toprak, E., & Gelbal, S. (2020). Comparison of classification performances of mathematics achievement at PISA 2012 with the artificial neural network, decision trees and discriminant analysis. *IJATE*, 7(4), 773-799.
- Toprak, E., & Kalkan, Ö. K. (2019a, June). Ergenlerde bilişsel esnekliğin yordayıcısı olarak başa çıkma stratejileri ve yetkinlik inançları [*Coping strategies and self-efficacy beliefs as predictors of cognitive flexibility in adolescents*]. 6. EJER Congress, Ankara.
- Toprak, E., & Kalkan, Ö. K. (2019b, Haziran). *Matematik başarısını yordayan değişkenlerin yapay sinir ağları ile incelenmesi* [*Examining the variables that predict mathematics achievement with artificial neural networks*]. 2. IBAD Congress, İstanbul.
- Turhan, K., Buirçin, K., & Engin, Y. Z. (2013). Estimation of student success with artificial neural networks. *Education and Science*, 38(170), 112-120.
- Yüzük, F. (2019). *Multiple regression analysis and neural networks with Turkish energy demand forecast* (Unpublished master's thesis). Cumhuriyet University.
- Zaidah, I., & Daliela, R. (2007). *Predicting students' academic performance: Comparing artificial neural network, decision tree and linear regression*. 21<sup>st</sup> Annual SAS Malaysia Forum, Kuala Lumpur.

**Información de contacto:** Emre Toprak. Erciyes University, Faculty of Education, Department of Educational Sciences. Postal code 38039 Kayseri, Turkey. E-mail: [etoprak@erciyes.edu.tr](mailto:etoprak@erciyes.edu.tr)