

Metodología de Análisis Envolvente de Datos (DEA) - GLMNET para la Evaluación y Pronóstico de Eficiencia Financiera en una Zona Franca Industrial - Colombia.

Tomás J. Fontalvo^{(1)*}, Enrique J. De la Hoz⁽²⁾ y Saúl Olivos⁽¹⁾

(1) Facultad de Ingeniería, Programa de Ingeniería Industrial, Universidad Libre, Barranquilla
(e-mail: tfontalvo@unilibrebaq.edu.co; solivo@unilibrebaq.edu.co)

(2) Facultad de Ingeniería, Programa de Ingeniería Industrial, Univ. Tecnológica de Bolívar, Campus Tecnológico, Cartagena, Colombia (e-mail: edelahoz@utb.edu.co)

*Autor a quien debe ser dirigida la correspondencia

Recibido Dic. 12, 2018; Aceptado Feb. 21, 2019; Versión final May. 5, 2019, Publicado Oct. 2019

Resumen

Se propone una metodología de evaluación y pronóstico para las empresas de la Zona Industrial del Puerto de la ciudad de Barranquilla - Colombia. Basado en un análisis empírico y racional, soportado en los conceptos de la eficiencia técnica, puramente técnica, la eficiencia aditiva, eficiencia de escala y de mezcla, así como en el algoritmo de Aprendizaje Automático GLMNET. Se trabajó con 29 empresas que presentaron sus estados financieros del año 2017 en la Cámara de Comercio de Barranquilla. Como resultado se encontró una eficiencia técnica promedio de 72.79%, una eficiencia puramente técnica de 82,54% y una eficiencia aditiva de 59,45 %. Adicionalmente se aportan las proyecciones requeridas para lograr que las organizaciones ineficientes alcancen la eficiencia. Del estudio también se puede observar que 11 empresas se constituyeron en referentes evaluativos para medir las empresas de la Zona franca del Puerto de Barranquilla. El algoritmo GLMNET arrojó un buen resultado en la clasificación de empresas eficientes y no eficientes del 93.1% de precisión.

Palabras clave: DEA; GLMNET; eficiencia; aprendizaje automático; zona franca

Methodology of Data Envelopment Analysis (DEA) - GLMNET for Assessment and Forecasting of Financial Efficiency in a Free Trade Zone - Colombia

Abstract

This research work proposes a methodology for evaluation and forecasting for companies located in the Industrial Port Zone of the city of Barranquilla, Colombia. Based on an empirical and rational analysis, supported by the concepts of technical efficiency, purely technical efficiency, additive efficiency, efficiency of scale and of mixing, as well as in the algorithm for machine learning GLMNET. Work was done with 29 companies that presented their complete financial statements for the year 2017 in the Chamber of Commerce of Barranquilla - Colombia. As a result, it was found an average technical efficiency of 72.79%, a purely technical efficiency of 82.54% and an additive efficiency of 59.45%. In addition, the projections required to make inefficient organizations achieve efficiency are contributed. From the study, it can also be observed that 11 companies were constituted as benchmarks to measure the companies of the Free Zone of the Port of Barranquilla. It is noteworthy that the algorithm GLMNET delivered a good result in the prediction of group membership of efficient and inefficient enterprises, with an accuracy of 93.1%.

Keywords: DEA; GLMNET; efficiency; machine learning; free trade zone

INTRODUCCIÓN

Actualmente las Zonas Francas en el mundo juegan un papel importante en el desarrollo y consolidación económica de una región, país o ciudad, esto las ha convertido en objeto de estudio y análisis como se aprecia en (Calix, Vigier, y Briozzo, 2015; De La Hoz, González, y Santana, 2016; Escandon-Barbosa y Hurtado-Ayala, 2016). Desde esta perspectiva, esta investigación cobra importancia al analizar el comportamiento de los diferentes tipos de eficiencia financiera de las empresas emplazadas en la zona franca industrial de Barranquilla, siendo la Zona Industrial del puerto, la más grande de la ciudad y una de las más importantes de la región por su impacto en la generación de empleo y divisas. En concordancia con esto, el desarrollo de este estudio es de vital importancia empresarial, pues a través de la medición de los tipos de eficiencia a estudiar se puede analizar el nivel de eficiencia de cada empresa, con lo que se aporta a la comunidad en general y en especial a los clientes y proveedores las características en términos de eficiencia de las empresas objeto de estudio. Adicionalmente se mide el desempeño y particularmente la eficiencia financiera del grupo empresarial a través de 5 tipos de eficiencias, con lo que evalúan integralmente las empresas objeto de esta investigación. Lo anterior es particularmente útil para entidades reguladoras y agencias de promoción de inversiones, dado que les permite contar con información objetiva para determinar planes de acción y mejoramiento, además de hacer un uso eficiente para la asignación de los recursos de apalancamiento productivo. Sin embargo, algo sustancialmente importante de esta investigación, es poder analizar cinco tipos de eficiencias financieras que permitan posteriormente por medio del algoritmo de Machine Learning, predecir si una organización es eficiente o no. Análisis decisivo y clave para todas las empresas o grupos de interés, para la toma de decisiones en la selección u alianza con estas empresas, clasificadas como eficientes o no eficientes. De igual manera el objeto de estudio de esta investigación al ser en una Zona Franca Industrial se constituye en un referente internacional para el análisis y toma de decisiones en este tipo de contexto a nivel regional e internacional.

Teniendo en cuenta que las exportaciones generan un incremento de la productividad como lo señalan diferentes autores (Pardo y García, 1999), esta variable se estudia en esta investigación toda vez que se aportan las directrices para que las empresas ineficientes de la zona franca en estudio alcancen la eficiencia empresarial. En términos financieros con esta investigación se analiza las magnitudes en términos de ingresos y utilidad neta de tal forma que las empresas ineficientes incrementen su eficiencia frente a los recursos que utilizan. Lo anterior permite generar una serie de decisiones en el contexto analizado que permite dinamizar y direccionar las acciones para alcanzar los resultados que permita a las organizaciones ineficientes ser competitivas en el sector. Existen investigaciones previas sobre el uso del análisis DEA y el aprendizaje automático, en (Hong et al., 1999) clasifican empresas eficientes y no eficientes utilizando DEA para luego predecir el estado de eficiencia de nuevas empresas, por otro lado (Lin, Hu, y Tsai, 2012) revisan en un periodo de cinco años la implementación del aprendizaje automático y los modelos DEA para la modelación del riesgo de quiebra en las empresas, a su vez otras investigaciones realizan procesos de clasificación paralelas utilizando la técnica de Support Vector Machines (Yeh, Chi, y Hsu, 2010), además de valoraciones empresariales en sectores específicos, como cadenas de suministro (Wong y Wong, 2007) y desempeño corporativo (Mirhedayatian, Azadi y Saen, 2014).

Por lo anterior en esta investigación se evalúa el sector empresarial mencionado apoyándose en conceptos de Análisis envolvente de datos DEA y aprendizaje automático. Con el fin de evaluar y pronosticar la eficiencia radial, la no radial y la aditiva de las unidades empresariales objeto de estudio. De lo anterior surgen las siguientes preguntas problemas de esta investigación ¿Cuál sería la estructura de las variables de entrada y las variables de salida para calcular las eficiencias técnicas, las eficiencias puramente técnicas y la eficiencia radial? ¿Cómo determinar la eficiencia de escala y la eficiencia de mezcla de las organizaciones objeto de esta investigación? ¿Qué modelo y estructura de aprendizaje automático permite establecer la pertenencia de las empresas eficientes y no eficientes de la zona franca de puerto de la ciudad de Barranquilla? A pesar de que existen diferentes investigaciones realizadas en áreas y saberes parecidos. Lo significativamente importante de esta investigación y el aporte al conocimiento, es el hecho de plantear una metodología articulando DEA y una técnica de machine Learning para el estudio, evaluación y pronóstico de una población “pequeña” como lo es la Zona Franca objeto de esta investigación.

Análisis Envolvente de datos DEA

El análisis Envolvente de datos es una técnica que permite analizar diferentes unidades tomadoras de decisión, que para nuestra investigación son las empresas objeto de estudio de la Zona Franca, creando un límite de eficiencia para su análisis y definición. El autor (Fontalvo, 2017) ha aplicado los modelos DEA y a su vez contextualizado el desarrollo de Modelos CCR-O, BBC-O y SBM-O que permiten evaluar la eficiencia radial, no radial y aditiva y a su vez estas eficiencias permiten realizar los cálculos de eficiencia de mezcla y escala, las cuales se evalúan en esta investigación.

Es importante señalar que estudio recientes muestran la pertinencia e importancia de la utilización del análisis envolvente de datos (Koçak, 2011), al desarrollar un sistema de indicadores para evaluar la eficiencia de aeropuertos en Turquía, por su parte (Wong y Wong, 2007) analizan diferentes cadenas de suministros bajo parámetros de eficiencia, lo cual muestra el uso de las técnicas desarrolladas en esta investigación aplicadas a diferentes contextos empresariales. De igual manera otros investigadores han utilizado y resaltado el uso de diferentes métodos de valoración de eficiencia en contextos de toma de decisiones complejas como en salud pública, gestión de primas de riesgo y análisis de datos estudiantiles (Falavigna et al., 2018; Galagedera et al., 2018; Zhang et al., 2017).

Algoritmo GLMNET

La técnica de aprendizaje automático GLMNET es un algoritmo de aprendizaje no supervisado del tipo ensamble (McCullagh, 1984), el cual desarrolla un modelo lineal generalizado conjunto (McCullagh, 1984), en el cual se enlazan el modelo de regresión de Ridge (Hoerl y Kennard, 1970) y la regresión de Lasso (Hans, 2009). El proceso de enlace se realiza a través de la penalización de la magnitud y el número de coeficientes del modelo final de regresión, de esta forma se evita el problema del sobre ajuste. Este tipo de modelos ha sido evaluado como efectivos en conjuntos de datos con muchas variables predictoras y pocas observaciones individuales. Así, en el caso límite cuando el parámetro de enlace toma un valor de cero, el modelo se desarrollará puramente como una regresión de Ridge y cuando toma un valor de 1 se comportará como una regresión de Lasso pura, en otra situación, por ejemplo, un valor de $\lambda = 0.05$ dará como resultado un modelo con un 95% de regresión de Ridge y un 5% de regresión de Lasso. Esta penalización es particularmente útil para conjuntos de datos que cuentan con pocas observaciones o cuando el número de variables predictoras es grande.

El algoritmo GLMNET desarrollado en (Friedman et al., 2010), define una variable de respuesta $Y \in \mathbb{R}$ un vector de variables predictoras $X \in \mathbb{R}^p$ aproximando las funciones de regresión por el modelo lineal $E(Y|X=x) = \beta_0 + X^T \beta$ así se cuenta con N parejas de observaciones (X_i, Y_i) , resolviendo el siguiente problema de optimización sin restricciones.

$$\min_{(\beta_0, \beta) \in \mathbb{R}^{p+1}} \left[\frac{1}{2N} \sum_{i=1}^N (Y_i - \beta_0 - x_i^T \beta)^2 + \lambda P_\alpha(\beta) \right] \quad (1)$$

Lo anterior es el sustento para el proceso de pronóstico de las empresas en el objeto de estudio de esta investigación.

METODOLOGÍA

Esta es una investigación de tipo evaluativa. Para el desarrollo de esta investigación se partió de un análisis racional que permitió definir las variables de entrada y salida para la aplicación de la técnica de Análisis Envolvente de Datos, asimismo se realizó un análisis empírico, sustentado en la información generada por cada unidad empresarial apoyada en la técnica de análisis envolvente de datos, particularmente tres modelos DEA CCR-O, BBC-O y SBM-O-C, que posibilitaron el cálculo de las diferentes eficiencias. Para lo anterior se partió de la información primaria aportada por las organizaciones de la Zona Franca del estudio a la Cámara de comercio de Barranquilla Colombia, seguidamente se procedió a analizar y a seleccionar las empresas que presentaron su información contable de forma completa. Para la selección de las variables a analizar en esta investigación se revisaron otras investigaciones que han seleccionados, estos rubros o rubros parecidos para realizar análisis de eficiencia y procesos de clasificación con otras técnica y enfoques (Fontalvo et al., 2018; González y Bermúdez, 2018). De donde se seleccionaron de forma racional propositiva los rubros que eran consistentes para la aplicación de la metodología de evaluación y pronóstico. Las variables de entrada usadas son, Total Activo, Patrimonio Neto, total pasivo, Activos corrientes, activos fijos gastos operaciones. Como variables de salida se utilizaron la utilidad bruta e ingresos de la operación. En la Figura 1, se ilustra las etapas del proceso de articulación del análisis de eficiencia y la predicción de las empresas.

Con las variables de entradas asociadas a los rubros financieros, se calcularon la eficiencia técnica, la eficiencia puramente técnica y las eficiencias aditivas, permitiendo así determinar las eficiencias de escala y la eficiencia de mezcla. Para lo anterior, se realizaron los cálculos con el software R (Team, 2013), y las técnicas de análisis envolvente de datos DEA (Simm y Besstremyannaya, 2015) y el paquete GLMNET (Friedman et al., 2010), siendo este el modelo de aprendizaje automático más pertinente para el presente estudio, por su buen desempeño y capacidad de clasificación para conjuntos de datos que cuentan con pocas observaciones.

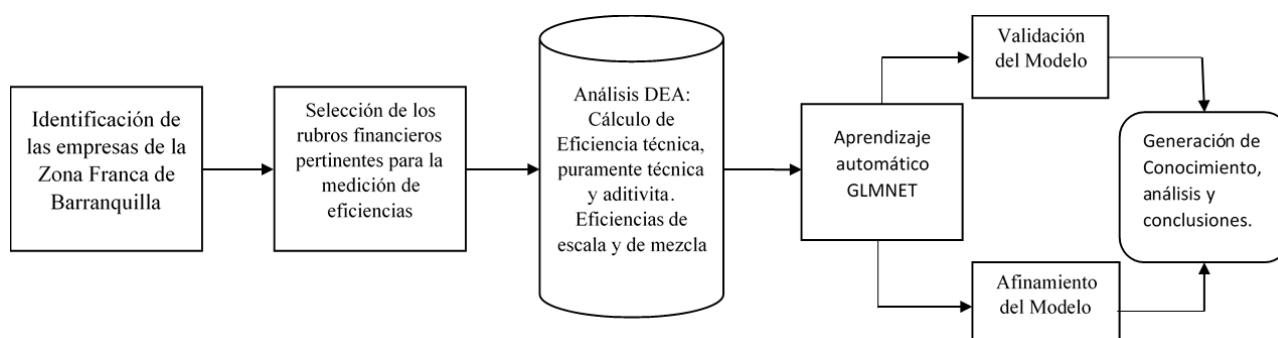


Fig. 1: Metodología de evaluación y pronóstico de la eficiencia financiera

RESULTADOS Y DISCUSIÓN

A partir de los rubros financieros de las organizaciones objeto de esta investigación, presentados en la metodología de la investigación se calcularon las eficiencias técnicas CCR-O y la puramente técnica BBC. De la Tabla 1 se puede observar que las primeras once empresas objeto de esta investigación son eficientes totalmente a alcanzar el 100 % en las eficiencias de técnica, puramente técnica y eficiencia aditiva, en la eficiencia de mezcla y escala, lo cual se genera al incrementar la producción en las empresas de la Zona Franca. Acerca de la eficiencia de escala se puede conceptuar que la producción de estas es sensible a la mejora para incrementar la eficiencia empresarial. Es importante señalar que cuando se revisa la eficiencia puramente técnica o BBC-O el grupo empresarial tiene un mejor desempeño, esta eficiencia no radial muestra que individualmente las empresas estudiadas tienen un buen rendimiento técnico o eficiencia empresarial. Lo que no ocurre cuando la medición se realiza con el modelo CCR-O, es decir cuando el grupo empresarial se compara con la eficiencia radial.

Tabla 1: Ranking de los diferentes tipos de eficiencias de las empresas de la zona Franca del Puerto de Barranquilla

	DMU	CCR-O	BBC-O	SBM-O	Eficiencia De Escala	Eficiencia De Mezcla
1	Actividades portuarias del caribe s.a.s.	1	1	1	1	1
2	Comercial de negocios zona franca s.a.	1	1	1	1	1
3	Gestion y gerencia logistica sociedad anonima simplificada	1	1	1	1	1
4	Barranquilla industrial de confecciones s.a.	1	1	1	1	1
5	Metalequipo y reciclados sas	1	1	1	1	1
6	Siscom zona franca s.a.	1	1	1	1	1
7	Solmico oil s.a.s.	1	1	1	1	1
8	Terminal de liquidos de barranquilla s.a.s.	1	1	1	1	1
9	Transcont s.a.s.	1	1	1	1	1
10	Z.f. construcciones s.a.	1	1	1	1	1
11	Quimpac logistica s.a.s.	1	1	1	1	1
12	Adama andina b.v. sucursal Colombia	0,87946	1	0,68055	0,87946	0,77382
13	Laboratorios rety de colombia s.a.s.	0,80245	1	6,59E-04	0,80245	0,00082
14	Plataforma petrolera del rio magdalena petromag s.a.s.	0,72863	1	0,67881	0,72863	0,93163
15	Industria articueros s.a.	0,70944	1	0,17488	0,70944	0,24651
16	Compañía envasadora del atlantico s.a.s.	0,66481	1	0,57421	0,66480	0,86373
17	Superbrix s.a.	0,54152	1	0,46748	0,54152	0,86328
18	Bero industrial s.a.s.	0,7751	0,92727	0,70799	0,83596	0,91334
19	Logistica e inversiones global	0,76693	0,78738	0,76275	0,97321	0,99538
20	Modatex y puntadas s.a.s.	0,69572	0,78687	0,45141	0,88416	0,64883

Tabla 1 (continuación)

21	Kruger park enterprises s.a.s.	0,61076	0,61237	0,48261	0,99735	0,79018
22	Colombian standard logistic s.a.	0,58825	0,64134	0,14802	0,91721	0,25163
23	Ensambladora sigma motors s.a.s. sigla e.s.m. s.a.s.	0,58272	0,60446	0,34085	0,96404	0,58492
24	Sedas del caribe s.a.s.	0,54844	0,59566	5,95E-02	0,92072	0,10842
25	Zona franca industrial de bienes y de servicios de barranquilla	0,41264	0,90001	0,39280	0,45848	0,95192
26	Mademeco s.a.	0,26908	0,39716	2,50E-03	0,67751	0,00930
27	Casa santana ron y licores s.a.s.	0,22566	0,29985	0,19886	0,75257	0,88125
28	Compania colombiana de ingenieria metalmeccanica s.a. sigla	0,19267	0,24764	2,63E-03	0,77802	0,01362
29	Maquiplast s.a.s.	0,11576	0,13891	0,11568	0,83337	0,99930

De igual forma en la Tabla 1 se observa la eficiencia Aditiva que tiene en cuenta la eficiencia radial y no radial de grupo de DMU analizadas. De la Tabla 2 se puede observar que esta eficiencia es la de menor desempeño, considerando que este modelo considera los rendimientos de escala. Por lo anterior para esta investigación se tomó este modelo aplicado para establecer las proyecciones requeridas para lograr que las empresas ineficientes alcancen su eficiencia como se puede observar en la Tabla 3.

Tabla 2: Promedio de los diferentes tipos de eficiencia

Modelo DEA utilizado	CCR-O	BBC-O	SBM-O-C
Promedio de eficiencia	0,72791	0,82548	0,59455
Desviación estándar	0,27521	0,26146	0,37691
Máximo	1	1	1
Mínimo	0,11576	0,13890	0,0006

Es importante desatacar que en otros estudios similares en empresas con las mismas características en el mismo contexto, los resultados de eficiencia muestran desempeños inferiores cuando se comparan con los obtenidos en el sector objeto de esta investigación (Fontalvo et al., 2014) donde obtuvo promedio de eficiencia del 56,62%, Por otro lado otros estudios similares en empresas que también desarrollan actividades de importación y exportación (Fontalvo et al., 2015) también se evidencian desempeños muy inferiores a los encontrados en la presente investigación. Pues solo se alcanzan un 31,85% de empresas eficientes. Lo que permite aseverar que existe evidencia empírica para afirmar que las empresas de la zona Franca estudiada analizadas tienen un buen desempeño cuando se comparan con empresas de contexto similares.

De la Tabla 3 se puede observar que de las 29 empresas objeto de esta investigación, 18 mostraron ser ineficientes. Para alcanzar y mejorar su eficiencia empresarial se requiere que los Ingresos de la operación y la utilidad se incrementen como se señala en la Tabla 3.

Tabla 3: Proyección de los rubros para alcanzar la eficiencia por parte de las empresas ineficientes

	DMU	Eficiencia SBM-O-C	Proyección de los Ingresos	Proyección de la Utilidad Neta
1	Adama andina b.v. sucursal Colombia	0,68055	0	2,09E+10
2	Bero industrial s.a.s. sigla bero industrial s.a.s.	0,70799	3,52E+08	1,15E+08
3	Casa santana ron y licores s.a.s.	0,19886	1,3E+10	3,36E+09
4	Colombian standard logistic s.a.	0,14802	7,3E+09	2,64E+08
5	Compania colombiana de ingenieria metalmeccanica s.a. sigla	2,63E-03	1,04E+10	5,66E+09
6	Compañía envasadora del atlantico s.a.s.	0,57421	6,74E+10	1,19E+10
7	Ensambladora sigma motors s.a.s. sigla e.s.m. s.a.s.	0,34085	7,07E+08	3,7E+08
8	Industria articueros s.a.	0,17488	3,82E+09	7,82E+09
9	Kruger park enterprises s.a.s.	0,48261	6,47E+08	1,18E+09
10	Laboratorios rety de colombia s.a.s. sigla retycol s.a.s.	6,59E-0	0	2,38E+10

Tabla 3 (continuación)

11	Logística e inversiones global s.a.s sigla inglobal s.a.s	0,76278	1,08E+09	2,01E+08
12	Mademeco s.a.	2,50E-3	8,18E+09	5,91E+09
13	Maquiplast s.a.s.	0,11568	1,67E+09	1,31E+08
14	Modatex y puntadas s.a.s.	0,45141	4,95E+08	2,97E+08
15	Plataforma petrolera del río Magdalena Petromag s.a.s.	0,67881	3,65E+09	4,31E+09
16	Sedas del Caribe s.a.s.	5,95E-02	3,81E+09	3,63E+09
17	Superbrix s.a.	0,46748	1,94E+10	1,35E+09
18	Zona Franca Industrial de Bienes y de Servicios de Barranquilla	0,392805	1,28E+11	1,72E+10

De la Tabla 4 se pueden observar las organizaciones que se constituyeron en referentes para evaluar a las demás organizaciones, esta frecuencia de referencia es de vital importancia para determinar las oportunidades de mejoramiento de una empresa en particular y determinar organizaciones modelos a seguir. Las mejores empresas del sector fueron las siguientes entidades.

Tabla 4: Empresas que se constituyeron en referentes para medir otras empresas

Nombre de la Organización (DMU)	Frecuencia de referencia de evaluación
Actividades portuarias del Caribe s.a.s. sigla Acticaribe s.a.s.	0
Barranquilla Industrial de Confecciones s.a.	2
Comercial de Negocios Zona Franca s.a.	0
Gestión y Gerencia Logística Sociedad Anónima Simplificada	1
Metalequ coast y reciclados sas	6
Quimpac Logística s.a.s.	1
Siscom Zona Franca s.a.	18
Solmico Oil s.a.s.	0
Terminal de Líquidos de Barranquilla s.a.s. sigla Telba s.a.s.	6
Transcont s.a.s.	4
Z.f. Construcciones s.a.	9

Modelo de clasificación GLMNET

Con base en la clasificación de empresas eficientes y no eficientes se procedió a utilizar el modelo de Machine Learning GLMNET, con lo que se analizó el proceso de pronóstico de la pertenencia de empresas a una de estas dos categorías, 1) Empresas eficientes y 2) Empresas no eficientes. Es importante señalar que otros autores han utilizado modelos similares asociados al cálculo multivariante para desarrollar procesos de pronóstico y pertinencia en sectores similares (Fontalvo et al., 2014). En estos, los niveles de determinación de pertenencia son menores a los alcanzados en esta investigación. Cabe señalar que otros estudios, muestran la importancia y la utilización de criterios de medición del desempeño, así como utilizar estructuras asociadas a la naturaleza propia de estas empresas para mejorar estos niveles de desempeño y eficiencia del tipo de organizaciones objeto de esta investigación (Fontalvo et al., 2018).

Con base en los resultados del proceso de clasificación empresarial en grupos de empresas eficientes y no eficientes, se estructuró un método de evaluación y pronóstico utilizando la metodología DEA - GLMNET. Para el desarrollo computacional se utilizó el software R (Team, 2013) y el paquete GLMNET (Friedman et al., 2010), en el cual se realizó una regresión ponderada con parámetros $\alpha = 1$ y $\lambda = 0.004$, generando así un modelo de regresión de ridge al 99.6% y de Lasso al 0.4%. Como lineamiento de calibración y aprendizaje del modelo se utilizó la técnica de validación cruzada, aprovechando el máximo de cada vector de datos para un aprendizaje y entrenamiento robusto del modelo, creando para esto 10 nuevos conjuntos de datos diferentes a través de un proceso de muestreo consecutivo. Para lo anterior, se han tomado como variables predictoras los rubros financieros indicados anteriormente y como variable de predicción la clasificación de empresas eficientes y no eficientes resultantes de la fase de aplicación del Análisis Envolvente de Datos (DEA).

El modelo de Aprendizaje Automático basado en el algoritmo de aprendizaje no supervisado GLMNET, entregó una precisión en el pronóstico de la pertenencia a los grupos de empresas eficientes y no eficientes del 93.1%, lo que muestra la alta capacidad de la metodología propuesta para clasificar grupos empresariales como eficientes y no eficientes. En la Tabla 5, se muestra la matriz de confusión resultante del proceso de clasificación empresarial, donde se evidencia la precisión y capacidad de clasificación del modelo propuesto.

Tabla 5: Matriz de confusión del proceso de clasificación y pronóstico

		Precisión 93.1 %	
		Realidad	
		Grupo	
Predicción	No eficiente	62.1%	6.9%
	Eficiente	0%	31%

En los modelos de aprendizaje automático es relevante analizar además de la precisión en la clasificación, el poder de discriminación por categorías, para este caso eficientes y no eficientes. Así para el modelo de clasificación de empresas eficientes y no eficientes se obtiene un valor de sensibilidad del 100% y de especificidad del 82%. Por ejemplo, encontramos que el modelo es capaz de identificar todas las empresas No eficientes y para las empresas Eficientes acierta en el 82% de los casos. La métrica del área bajo la curva para analizar los valores de sensibilidad y especificidad conjuntamente, arroja un valor del 90%, evidenciando que la alta capacidad del modelo para identificar los patrones de comportamiento financiero que diferencian las empresas eficientes de las no eficientes no se debe al azar, sino a la sincronización del método propuesto, generando así una información pertinente y robusta, claramente alejada del comportamiento de un método aleatorio de clasificación. Lo anterior es coherente con otros estudios similares en donde utilizan otras técnicas de aprendizaje automático (Fontalvo y De la Hoz., 2018; Fontalvo et al., 2018; De la Hoz et al., 2019) que permiten identificar patrones de comportamiento en un contexto particular.

Como proyección de otras investigaciones se invita los investigadores a proponer modelos, técnicas y herramientas que involucren otras variables para medir la eficiencia y otros modelos de aprendizaje automático específicos para predecir y pronosticar con el fin analizar el incremento de la precisión para predecir la pertenencia de empresas eficientes o no en las zonas Francas de cualquier contexto nacional o internacional.

CONCLUSIONES

Como resultado de esta investigación se pueden precisar las siguientes conclusiones y aportes generados 1). La estructura para evaluar la eficiencia radial, no radial y aditiva, así como la de escala y mezcla financiera de empresas de una zona franca desde diferentes modelos, y adicionalmente se establece una relación de los rubros financieros Total Activo, Patrimonio Neto, Total Pasivo, Activos Corrientes, Activos Fijos y Gastos Operaciones asociados a la Utilidad Bruta y los Ingresos de la Operación y su relación con la eficiencia técnica o administrativa, la eficiencia puramente técnica y la eficiencia aditiva. 2) Se establece un análisis de causalidad entre una serie de rubros financieros con los diferentes tipos de eficiencias. Adicionalmente se analiza los rubros o variables del modelo DEA y su incidencia con la eficiencia de escala. 3) Una metodología cuantitativa que articula la técnica de Análisis envolvente de datos DEA con un modelo de Aprendizaje Automático, lo cual se generó producto de un análisis racional propositivo que estructuró y contrastó los resultados de la técnica de análisis envolvente de datos con las entradas para el modelo de aprendizaje automático GLMNET. 4) Una metodología replicable y reproducible, que puede ser aplicada en otros entornos industriales de comercio exterior, siempre que se cuente con los rubros financieros aquí estudiados.

De igual forma se puede concluir la pertinencia y oportunidad del uso del algoritmo de Machine Learnig Glnet, para el proceso de evaluación y pronóstico de poblaciones pequeñas.

REFERENCIAS

- Calix, C., H. Vigier y A. Briozzo, Capital intelectual y otros determinantes de la ventaja competitiva en empresas exportadoras de la zona norte de Honduras, *Suma de Negocios*, 6(14), 130-137 (2015)
- De La Hoz, E., A. González y A. Santana, Metodología de Medición del Potencial Exportador de las Organizaciones Empresariales, doi:10.4067/S0718-07642016000600003, *Información Tecnológica*, 27(6), 11-18 (2016)

- De La Hoz, E., E. De la Hoz y T. Fontalvo, Metodología de Aprendizaje Automático para la Clasificación y Predicción de Usuarios en Ambientes Virtuales de Educación, doi:10.4067/S0718-07642019000100247, *Información Tecnológica*, 30 (1), 247-254 (2019)
- Escandon, D.M. y A. Hurtado, Influencia de los estilos de liderazgo en el desempeño de las empresas exportadoras coombianas, *Estudios Gerenciales*, 32(139), 137-145 (2016)
- Falavigna, G., R. Ippoliti y G. Ramello, DEA-based Malmquist productivity indexes for understanding courts reform, *Socio-Economic Planning Sciences*, 62, 31–43 (2018)
- Fontalvo, T.J., E. De la Hoz y E. De La Hoz, Método Análisis Envolvente de Datos y Redes Neuronales en la Evaluación y Predicción de la Eficiencia Técnica de Pequeñas Empresas Exportadoras, doi:10.4067/S0718-07642018000600267, *Información Tecnológica*, 29(6), 267- 275 (2018)
- Fontalvo, T. J., Eficiencia de las entidades prestadoras de salud (EPS) en Colombia por medio de análisis envolvente de datos, *Ingeniare Revista chilena de ingeniería*, 25(4), 681–692 (2017)
- Fontalvo, T.J., E. De la Hoz y J. Morelos, Aplicación del análisis discriminante para evaluar el mejoramiento de los indicadores financieros en las empresas del sector extracción de petróleo crudo y gas natural en Colombia, *Revista Soluciones de Postgrado EIA*, 1(7), 1-16 (2014)
- Fontalvo, T.J., E. De la Hoz. A. Mendoza, Aplicación de Minería de Datos para la Clasificación de Programas Universitarios de Ingeniería Industrial Acreditados en Alta Calidad en Colombia, doi:10.4067/S0718-07642018000300089, *Información Tecnológica*, 29(3), 89-96 (2018)
- Fontalvo, T. J. y E. De La Hoz, Study of financial efficiency in companies certified with the BASC label using Data Envelopment Analysis: Case applied in Cali – Colombia, *Entramado*, 14(1), 78-87 (2018)
- Fontalvo, T. J., A. Mendoza y D. Visbal, Efficiency in logistics processes in Medellin BASC certified companies through data envelopment analysis, *Revista U.D.C.A Actualidad & Divulgación Científica*, 17(1), 265-274 (2014)
- Fontalvo, T. J., A. Mendoza y D. Visbal, Análisis comparativo de eficiencia financiera: Estudio de un caso sectorial en Barranquilla, *Prospectiva*, 13(2), 16–24 (2015)
- Friedman, J., T. Hastie y R. Tibshirani, Regularization paths for generalized linear models via coordinate descent, *Journal of statistical software*, 33(1), 1-22 (2010)
- Galagedera, D.U., I. Roshdi y otros dos autores, New network DEA model for mutual fund performance appraisal: An application to US equity mutual funds, *Omega*, 77(1), 168–179 (2018)
- Gonzalez, P. y T. Bermudez, Una Aproximación al Modelo de Toma de Decisiones Usado por los Gerentes de las Micro, Pequeñas y Medianas Empresas Ubicadas en Cali, Colombia desde un Enfoque de Modelos de Decisión e Indicadores Financieros y no Financieros, *Contaduría Unviersidad de Antioquia*, 52(1), 131-154 (2008)
- Hans, C., Bayesian lasso regression, *Biometrika*, 96(4), 835-845 (2009)
- Hoerl, A. E. y R. Kennard, Ridge regression: Biased estimation for nonorthogonal problems, *Technometrics*, 12(1), 55-67 (1970)
- Hong, H. K., S. Ha, y otros tres autores, Evaluating the efficiency of system integration projects using data envelopment analysis (DEA) and machine learning, doi:10.1016/S0957- 4174(98)00077-3, *Expert Systems with Applications*, 16(3), 283-296 (1999)
- Koçak, H., Efficiency examination of Turkish airport with DEA approach. *International Business Research*, 4(2), 204-212 (2011)
- Lin, W. Y., Y. Hu y C. Tsai, Machine Learning in Financial Crisis Prediction: A Survey, *IEEE Transactions on Systems, Man, and Cybernetics, Part C*, doi: 10.1109/TSMCC.2011.2170420, *Applications and Reviews*, 42(4), 421-436 (2012)
- McCullagh, P., Generalized linear models, *European Journal of Operational Research*, 16(3), 285-292 (1984)
- Mirhedayatian, S. M., M. Azadi y R. Saen, A novel network data envelopment analysis model for evaluating green supply chain management, doi: 10.1016/j.ijpe.2013.02.009, *International Journal of Production Economics*, 147, 544–554 (2014)
- Pardo, S.R. y J. García, Eficiencia empresarial y actividad exportadora: Diversificación y tamaño en las empresas industriales españolas, *Papeles de Economía Española*, 78, 220-235 (1999)
- Peng, W. y K. Yew, Supply chain performance measurement system using DEA modeling, doi: 10.1108/02635570710734271, *Industrial Management & Data Systems*, 107(3), 361-381 (2007)
- Simm, J. y G. Besstremyannaya, Package 'DEA', Version 1.2-4. August 4 (2015)
- Team, R. C., A language and environment for statistical computing (2013)
- Yeh, C., D. Chi y M. Hsu, A hybrid approach of DEA, rough set and support vector machines for business failure prediction, doi:10.1016/j.eswa.2009.06.088, *Expert Systems with Applications*, 37(2), 1535-1541 (2010)
- Zhang, Z., M. Wang y otros dos autores, Research on the development efficiency of regional high-end talent in China: A complex network approach, *PloS one*, 12(12), (2017)