



Paper Type: Original Article



An Integrated Machine Learning and Inverse Data Envelopment Analysis Framework for Predicting and Enhancing Bank Branch Efficiency

Akram Mohsenikia^{1,*}, Ghasem Tohidi¹, Masoud Sanei¹ , Behrooz Daneshian¹, Shabnam Rezavian²

¹ Department of Mathematics, Central Tehran Branch, Islamic Azad University, Tehran, Iran; Mohsenikia@gmail.com; Ghatohidi@yahoo.com; M_sanei@iauctb.ac.ir; Be_daneshian@iauctb.ac.ir.

² Department of Mathematics, South Tehran Branch, Islamic Azad University, Tehran, Iran; Sh_razavyan@azad.ac.ir.

Citation:



Mohsenikia, A., Tohidi, Gh., Sanaei, M., Daneshian, B., & Rezavian, Sh. (2026). An integrated machine learning and inverse data envelopment analysis framework for predicting and enhancing bank branch efficiency. *Strategic Studies in Financial Management and Insurance*, 3(3), 163-172.

Received: 04/02/2026

Reviewed: 05/04/2026

Revised: 20/04/2026

Accepted: 27/05/2026

Abstract

Purpose: This study aims to develop an intelligent decision-support framework for predicting and improving the efficiency of bank branches by integrating Data Envelopment Analysis (DEA), Machine Learning (ML), and Inverse Data Envelopment Analysis (Inverse DEA).

Methodology: First, the relative efficiency of bank branches was evaluated using DEA based on real operational data. Next, machine learning algorithms were employed to predict the future efficiency status of the branches. Finally, Inverse DEA was applied to determine the target input and output levels required to improve the performance of inefficient branches.

Findings: The results demonstrate that the proposed framework not only provides accurate efficiency assessment but also identifies the key factors affecting inefficiency, predicts future branch performance, and generates quantitative improvement targets. Furthermore, integrating ML with Inverse DEA improves prediction accuracy and managerial decision-making compared with using either method independently.

Originality/Value: The novelty of this study lies in proposing an integrated framework that combines efficiency evaluation, predictive analytics, and performance improvement within a unified decision-support system. The proposed approach offers a practical tool for optimal resource allocation and enhancing productivity across banking networks.

Keywords: Data envelopment analysis, Inverse data envelopment analysis, Machine learning, Intelligent decision support system, Bank branch efficiency.



Corresponding Author: Mohsenikia@gmail.com

<https://doi.org/10.22105/ssfmi.v3i3.115>

Licensee. **Strategic Studies in Financial Management and Insurance**. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0>).



چارچوب یکپارچه یادگیری ماشین و تحلیل پوششی داده‌های معکوس برای پیش‌بینی و ارتقای کارایی شعب بانک

اکرم محسنی کیا^۱، قاسم توحیدی^۱، مسعود صانعی^۱، بهروز دانشیان^۱، شبنم رضویان^۲

^۱گروه ریاضی، واحد تهران مرکز، دانشگاه آزاد اسلامی، تهران، ایران.

^۲گروه ریاضی، واحد تهران جنوب، دانشگاه آزاد اسلامی، تهران، ایران.

چکیده

هدف: این پژوهش با هدف ارایه یک چارچوب تصمیم‌یار هوشمند برای پیش‌بینی و ارتقای کارایی شعب بانک، از طریق ترکیب تحلیل پوششی داده‌ها، یادگیری ماشین و تحلیل پوششی داده‌های معکوس انجام شده است.

روش‌شناسی پژوهش: ابتدا کارایی نسبی شعب بانک با استفاده از تحلیل پوششی داده‌ها و بر اساس داده‌های عملیاتی واقعی ارزیابی شد. سپس الگوریتم‌های یادگیری ماشین برای پیش‌بینی وضعیت آتی کارایی شعب به‌کار گرفته شدند و در نهایت، تحلیل پوششی داده‌های معکوس برای تعیین مقادیر هدف ورودی‌ها و خروجی‌های شعب ناکارا به‌منظور بهبود عملکرد آن‌ها استفاده شد.

یافته‌ها: نتایج نشان داد چارچوب پیشنهادی علاوه بر ارزیابی دقیق کارایی، قادر به پیش‌بینی وضعیت آینده شعب، شناسایی عوامل موثر بر ناکارایی و ارایه اهداف کمی برای بهبود عملکرد است. همچنین، ترکیب یادگیری ماشین و تحلیل پوششی داده‌های معکوس، دقت پیش‌بینی و کیفیت تصمیم‌گیری مدیریتی را نسبت به استفاده مستقل از هر یک از این روش‌ها بهبود داد.

اصالت/ارزش افزوده علمی: نوآوری این پژوهش در ارایه یک چارچوب یکپارچه است که برای نخستین بار ارزیابی کارایی، پیش‌بینی هوشمند و تعیین راهکارهای بهبود عملکرد را در یک مدل تلفیق می‌کند. این چارچوب می‌تواند به‌عنوان ابزاری موثر برای تخصیص بهینه منابع و افزایش بهره‌وری در شبکه بانکی مورد استفاده قرار گیرد.

کلیدواژه‌ها: تحلیل پوششی داده‌ها، تحلیل پوششی داده‌های معکوس، یادگیری ماشین، سیستم تصمیم‌یار هوشمند، کارایی شعب بانک.

۱- مقدمه

در سال‌های اخیر رقابت در صنعت بانکداری به‌ویژه با توسعه بانکداری الکترونیک و فین‌تک‌ها به‌طور قابل‌توجهی افزایش یافته است. در چنین شرایطی، افزایش بهره‌وری و کارایی شعب بانک به یکی از مهم‌ترین دغدغه‌های مدیران شبکه بانکی تبدیل شده است. شعب بانک همچنان به‌عنوان هسته عملیاتی تجهیز منابع، تخصیص اعتبارات و ارایه خدمات مالی نقش اساسی در سودآوری بانک‌ها ایفا می‌کنند و عملکرد آن‌ها تاثیر مستقیمی بر رضایت مشتریان و سهم بازار دارد.

روش‌های سنتی ارزیابی عملکرد عموماً مبتنی بر شاخص‌های مالی منفرد هستند و توانایی بررسی همزمان چندین ورودی و خروجی را ندارند. با توجه به ماهیت چندبعدی عملکرد شعب بانک، استفاده از روش‌هایی که بتوانند همزمان منابع مصرفی و خدمات ارایه‌شده را تحلیل کنند ضروری

است. تحلیل پوششی داده‌ها^۱ نخستین بار توسط چارنز و همکاران [1] معرفی شد و به عنوان یک روش ناپارامتری مبتنی بر برنامه‌ریزی خطی^۲ برای اندازه‌گیری کارایی نسبی واحدهای تصمیم‌گیرنده^۳ توسعه یافت. در ادامه، مدل BCC^4 با فرض بازده متغیر نسبت به مقیاس توسط بنکر و همکاران [2] ارائه شد.

DEA طی چهار دهه گذشته به طور گسترده در ارزیابی کارایی بانک‌ها و موسسات مالی مورد استفاده قرار گرفته است [3-6]. با وجود مزایای DEA ، این روش صرفاً به اندازه‌گیری کارایی می‌پردازد و فاقد قابلیت پیش‌بینی وضعیت آینده است.

از سوی دیگر، الگوریتم‌های یادگیری ماشین^۵ قادر به کشف الگوهای پنهان داده‌ها و پیش‌بینی عملکرد واحدها هستند و در سال‌های اخیر کاربرد گسترده‌ای در تحلیل داده‌های مالی و بانکی یافته‌اند [7]، [8]. همچنین تحلیل پوششی داده‌های معکوس^۶ به عنوان ابزاری برای تعیین اهداف بهبود عملکرد و رسیدن به سطح کارایی مطلوب معرفی شده است [9].

هدف این پژوهش ارائه یک چارچوب یکپارچه برای ارزیابی، پیش‌بینی و بهبود عملکرد شعب بانک است. در این چارچوب ابتدا کارایی شعب با مدل BCC خروجی محور محاسبه می‌شود، سپس وضعیت کارایی با ML پیش‌بینی شده و در نهایت اهداف عملیاتی بهبود از طریق $Inverse DEA$ تعیین می‌گردد.

۲- مرور ادبیات پژوهش

ارزیابی کارایی بانک‌ها و موسسات مالی یکی از حوزه‌های پرکاربرد DEA محسوب می‌شود [3]. مطالعات متعددی از DEA برای اندازه‌گیری کارایی نسبی شعب بانک‌ها استفاده کرده‌اند [4-6]. تمرکز پژوهش‌ها بر رتبه‌بندی واحدها و تعیین شعب مرجع بود. با گذشت زمان، پژوهشگران دریافتند که اندازه‌گیری کارایی به تنهایی برای تصمیم‌گیری مدیریتی کافی نیست و مدیران علاوه بر دانستن وضعیت کارایی، نیازمند شناخت علل ناکارایی و مسیرهای بهبود عملکرد هستند.

$Inverse DEA$ به عنوان توسعه‌ای از DEA معرفی شد که قادر است مقادیر مورد نیاز ورودی‌ها یا خروجی‌ها را برای رسیدن به سطح کارایی مطلوب تعیین کند [9]. این روش به مدیران کمک می‌کند تا اهداف عملیاتی قابل اجرا برای واحدهای ناکارا مشخص نمایند. با این حال، $Inverse DEA$ نیز فاقد قابلیت پیش‌بینی و تحلیل علی بوده و تنها پس از مشاهده ناکارایی قابل استفاده است.

در سال‌های اخیر، کاربرد ML در تحلیل داده‌های مالی به طور قابل توجهی افزایش یافته است. الگوریتم‌هایی مانند جنگل تصادفی^۷، ماشین بردار پشتیبان^۸ و تقویت گرادیان^۹ توانایی بالایی در پیش‌بینی، طبقه‌بندی داده‌ها و تحلیل کارایی بانک‌ها دارند [7]، [8]، [10]. همچنین پژوهش‌های اخیر نشان می‌دهد که راهبردهای دیجیتال‌سازی خدمات بانکی و توان مدیریتی در تخصیص بهینه منابع، نقش معناداری در بهبود کارایی عملیاتی بانک‌ها ایفا می‌کنند [11]، [12].

ترکیب این روش‌ها با DEA امکان تبدیل یک ابزار اندازه‌گیری عملکرد به یک سیستم تصمیم‌یار هوشمند را فراهم می‌کند. در چنین چارچوبی، DEA وضعیت کارایی را تعیین می‌کند، ML عوامل ایجادکننده ناکارایی را شناسایی می‌نماید و $Inverse DEA$ برنامه بهبود عملکرد را ارائه می‌دهد.

¹ Data Envelopment Analysis (DEA)

² Linear Programming (LP)

³ Decision-Making Units (DMUs)

⁴ Banker, Charnes, and Cooper (BCC) model

⁵ Machine Learning (ML)

⁶ Inverse Data Envelopment Analysis (Inverse DEA)

⁷ Random Forest

⁸ Support Vector Machine (SVM)

⁹ Gradient Boosting (GB)

۳- روش تحقیق

در این پژوهش شعب بانک به‌عنوان *DMUs* در نظر گرفته شدند. داده‌ها شامل متغیرهای عملیاتی شعب در یک دوره زمانی مشخص است. ورودی‌ها شامل تعداد کارکنان، هزینه عملیاتی و حجم سپرده‌ها و خروجی‌ها شامل حجم تسهیلات اعطایی، کارمزد دریافتی و تعداد حساب‌های فعال در نظر گرفته شد.

به‌منظور سنجش کارایی نسبی شعب بانک از مدل *BCC* خروجی محور با فرض بازده متغیر نسبت به مقیاس استفاده شد [2]. انتخاب رویکرد خروجی محور به این دلیل است که مدیران شعب در کوتاه‌مدت کنترل محدودی بر منابع مصرفی داشته و تمرکز مدیریتی عمدتاً بر افزایش خدمات و بازدهی عملیاتی شعب از طریق افزایش تسهیلات اعطایی، کارمزدهای دریافتی و تعداد حساب‌های فعال است.

۳-۱- فرم‌بندی ریاضی مدل *BCC* خروجی محور

فرض کنید n شعبه با m ورودی و s خروجی وجود دارد. برای ارزیابی شعبه o یک مساله *LP* به‌صورت زیر تعریف می‌شود:

$$\begin{aligned} \max \quad & \phi_o, \\ \text{s.t.} \quad & \sum_{j=1}^n \lambda_j x_{ij} \leq x_{io}, \quad i=1, \dots, m, \\ & \sum_{j=1}^n \lambda_j y_{rj} \geq \phi_o y_{ro}, \quad r=1, \dots, s, \\ & \sum_{j=1}^n \lambda_j = 1, \\ & \lambda_j \geq 0, \quad j=1, \dots, n. \end{aligned} \quad (1)$$

x_{ij} مقدار ورودی i برای شعبه j و y_{rj} مقدار خروجی r برای شعبه j می‌باشد. ضرایب λ_j ترکیبی مجازی از شعب مرجع را تشکیل می‌دهند و مرز کارایی را برای مقایسه با شعبه موردبررسی می‌سازند.

در مدل خروجی محور *BCC*، مقدار بهینه ϕ_o^* نشان‌دهنده ضریب افزایش متناسب خروجی‌ها است و مستقیماً بیانگر کارایی نیست. در صورتی که $\phi_o^* = 1$ باشد، شعبه روی مرز کارایی قرار دارد؛ اما اگر $\phi_o^* < 1$ به‌دست آید، شعبه ناکارا بوده و لازم است خروجی‌های خود را حداقل به نسبت ϕ_o^* افزایش دهد تا به مرز کارایی منتقل شود. کارایی فنی برابر با معکوس این مقدار، به‌صورت زیر است:

$$TE_o = \frac{1}{\phi_o^*}. \quad (2)$$

۳-۲- مدل تحلیل پوششی داده‌های معکوس

مدل *Inverse DEA* به‌منظور تعیین اهداف عملیاتی بهبود عملکرد مورد استفاده قرار گرفت [9]. این مدل امکان تعیین حداقل تغییرات لازم در خروجی‌ها برای انتقال شعب ناکارا به مرز کارایی را فراهم می‌سازد. هدف *Inverse DEA* تعیین تغییرات موردنیاز در خروجی‌ها به‌گونه‌ای است که شعبه ناکارا بدون افزایش منابع مصرفی به مرز کارایی منتقل شود. در این پژوهش ورودی‌ها ثابت فرض شده و تنها مقادیر خروجی‌ها قابل بهبود در نظر گرفته می‌شوند.

فرض کنید شعبه o ناکارا باشد و مقادیر خروجی‌های آن برابر y_{ro} باشد. هدف، تعیین افزایش‌های حداقلی Δy_r برای خروجی‌ها است به‌طوری‌که شعبه به مجموعه کارا منتقل گردد.

مدل برنامه‌ریزی خطی *Inverse DEA* به‌صورت زیر فرموله می‌شود:

$$\begin{aligned}
 \min \quad & \sum_{r=1}^s w_r \Delta y_r, \\
 s.t. \quad & \sum_{j=1}^n \lambda_j x_{ij} \leq x_{io}, \quad i=1, \dots, m, \\
 & \sum_{j=1}^n \lambda_j y_{rj} \geq y_{ro} + \Delta y_r, \quad r=1, \dots, s, \\
 & \sum_{j=1}^n \lambda_j = 1, \\
 & \lambda_j \geq 0, \quad j=1, \dots, n, \\
 & \Delta y_r \geq 0, \quad r=1, \dots, s.
 \end{aligned} \tag{۳}$$

در این مدل، Δy_r بیانگر مقدار افزایش مورد نیاز در خروجی r است. جواب بهینه این مدل نشان می‌دهد هر شعبه دقیقاً در کدام خروجی (تسهیلات، کارمزد یا تعداد حساب‌ها) باید چه میزان بهبود داشته باشد تا کارا شود. بدین ترتیب *Inverse DEA* یک برنامه عملیاتی کمی برای مدیران شعب فراهم می‌کند.

۳-۳-۳ مدل یادگیری ماشین برای پیش‌بینی کارایی

پس از محاسبه امتیاز کارایی شعب با استفاده از مدل *BCC* خروجی محور، شعب به دو گروه کارا و ناکارا تقسیم شدند. در گام دوم چارچوب پیشنهادی، از یک مدل *ML* نظارت‌شده به منظور پیش‌بینی وضعیت کارایی شعب در دوره آتی استفاده شد. در این مرحله، *DEA* نقش تولیدکننده برچسب^۱ را ایفا می‌کند و امتیازهای کارایی حاصل از *DEA* مبنای تعریف متغیر وابسته در مدل *ML* قرار می‌گیرند. این رویکرد امکان انتقال از یک چارچوب صرفاً ارزیابانه به یک چارچوب پیش‌نگر را فراهم می‌سازد.

پس از محاسبه کارایی، این برچسب‌ها به عنوان متغیر هدف در مدل *ML* استفاده گردید. در این پژوهش از الگوریتم *Random Forest* که توسط بریمن [10] ارایه شد، استفاده گردید. این الگوریتم در مطالعات بانکی برای پیش‌بینی عملکرد و طبقه‌بندی کارایی کاربرد گسترده‌ای داشته است [7]، [8] و متغیرهایی مانند درجه شعبه، موقعیت مکانی، قدمت شعبه و تعداد دستگاه‌های خودپرداز به عنوان ویژگی‌های ورودی در مدل وارد شدند.

۳-۳-۱-۱ تعریف مساله یادگیری

در این پژوهش، هدف مدل *ML*، برآورد مقدار عددی کارایی نیست، بلکه پیش‌بینی وضعیت کارایی شعب در دوره آتی است. از این رو، مساله یادگیری به صورت یک مساله طبقه‌بندی دودویی^۲ تعریف شد که در آن برچسب‌ها از امتیازهای کارایی حاصل از *DEA* استخراج گردیدند.

برای هر شعبه o ، ابتدا کارایی فنی خروجی محور با نماد TE_o محاسبه شد. سپس متغیر هدف به صورت زیر تعریف گردید:

$$Z_o = \begin{cases} 1, & \text{if } TE_o = 1, & \text{Efficient Branch} \\ 0, & \text{if } TE_o < 1. & \text{Inefficient Branch} \end{cases} \tag{۴}$$

بنابراین متغیر وابسته مدل *ML* نشان‌دهنده وضعیت کارایی شعب است، نه مقدار کارایی. بردار ویژگی‌های هر شعبه در زمان t با X_o^t نمایش داده شد که هر مولفه آن نشان‌دهنده یکی از متغیرهای عملیاتی یا محیطی موثر بر عملکرد شعبه نظیر تعداد کارکنان، هزینه‌های عملیاتی، حجم سپرده‌ها، ویژگی‌های منطقه جغرافیایی و شاخص‌های تقاضای خدمات بانکی به صورت زیر است:

$$X_o^t = (x_{1o}^t, x_{2o}^t, \dots, x_{ko}^t). \tag{۵}$$

هدف مدل *ML*، برآورد تابع احتمال زیر است:

^۱ Label generator

^۲ Binary classification

$$P(Z_o^{t+1} = 1 | X_o^t), \quad (۶)$$

که احتمال کارا بودن شعبه در دوره آتی را بر اساس اطلاعات جاری آن تخمین می‌زند. این تعریف امکان شناسایی پیش‌دستانه شعب در معرض ناکارایی و مداخله مدیریتی قبل از افت عملکرد را فراهم می‌سازد و بدین ترتیب، مدل ML نقش بخش پیش‌نگر در سیستم تصمیم‌یار پیشنهادی را ایفا می‌کند.

۳-۳-۲- الگوریتم پیشنهادی جنگل تصادفی

به‌منظور مدل‌سازی رابطه میان ویژگی‌های عملیاتی شعب و وضعیت کارایی آن‌ها، از الگوریتم *Random Forest* استفاده شد. این الگوریتم یکی از روش‌های یادگیری تجمیعی^۱ مبتنی بر درخت‌های تصمیم است که با ترکیب مجموعه‌ای از طبقه‌بندهای ضعیف، یک مدل پیش‌بینی‌کننده پایدار و با قابلیت تعمیم بالا ایجاد می‌کند. انتخاب این روش به دلیل ماهیت پیچیده و غیرخطی رابطه میان منابع عملیاتی شعب و کارایی آن‌ها صورت گرفت؛ روابطی که معمولاً با مدل‌های خطی کلاسیک قابل استخراج نیستند.

فرض شود مجموعه داده آموزشی شامل N مشاهده باشد. در جنگل تصادفی ابتدا با استفاده از روش نمونه‌گیری بوت‌استرپ، تعداد B زیرمجموعه تصادفی از داده‌ها تولید می‌شود. هر زیرمجموعه برای آموزش یک درخت تصمیم مستقل مورد استفاده قرار می‌گیرد. در فرایند ساخت هر درخت، در هر گره به‌جای بررسی تمام ویژگی‌ها، تنها یک زیرمجموعه تصادفی از ویژگی‌ها با اندازه m_{inv} برای تعیین بهترین تقسیم انتخاب می‌شود. این دو سطح تصادفی‌سازی، یعنی تصادفی‌سازی در نمونه‌های آموزشی و تصادفی‌سازی در انتخاب ویژگی‌ها، موجب کاهش همبستگی میان درخت‌ها و در نتیجه کاهش واریانس پیش‌بینی می‌گردد.

پیش‌بینی نهایی مدل بر اساس رای‌گیری اکثریت درخت‌ها به‌دست می‌آید.

$$\hat{y}_{o,t+1} = \text{Mode}\{h_1(X_{o,t}), h_2(X_{o,t}), \dots, h_B(X_{o,t})\}. \quad (۷)$$

علاوه بر برچسب طبقه‌بندی، احتمال کارا بودن شعبه نیز از میانگین احتمالات پیش‌بینی‌شده توسط درخت‌ها محاسبه می‌شود.

$$\hat{P}(y_{o,t+1} = 1) = \frac{1}{B} \sum_{b=1}^B P_b(y_{o,t+1} = 1). \quad (۸)$$

به‌کارگیری جنگل تصادفی در این پژوهش چند مزیت مهم ایجاد می‌کند. نخست آنکه این الگوریتم قادر به مدل‌سازی روابط غیرخطی و اثرات متقابل میان متغیرهای عملیاتی نظیر تعداد کارکنان، سپرده‌ها و هزینه‌های عملیاتی است. دوم آنکه در برابر هم‌خطی متغیرها و داده‌های پرت که از ویژگی‌های رایج داده‌های بانکی محسوب می‌شوند، مقاومت بالایی دارد. افزون بر این، جنگل تصادفی امکان محاسبه شاخص اهمیت ویژگی‌ها^۲ را فراهم می‌سازد که از طریق آن می‌توان سهم هر متغیر در پیش‌بینی ناکارایی شعب را تحلیل کرد. این قابلیت، مدل ML را از یک ابزار صرفاً پیش‌بینی‌کننده به ابزاری تحلیلی برای پشتیبانی تصمیمات مدیریتی تبدیل می‌کند.

۳-۳-۳- فرایند آموزش و تنظیم مدل

پیش از آموزش مدل، داده‌ها از نظر مقادیر گمشده و ناهنجاری بررسی و در صورت لزوم پاک‌سازی شدند. سپس مجموعه داده به دو بخش آموزش (۷۰٪) و آزمون (۳۰٪) تقسیم گردید. به‌منظور جلوگیری از بیش‌برازش و افزایش قابلیت تعمیم، تنظیم ابرپارامترهای مدل شامل تعداد درخت‌ها (B)، حداکثر عمق درخت و تعداد ویژگی‌های انتخابی در هر گره با استفاده از جستجوی شبکه‌ای^۳ و اعتبارسنجی متقابل ده‌تایی^۴ انجام شد.

این رویکرد امکان انتخاب ترکیب بهینه پارامترها را بر اساس عملکرد میانگین مدل در داده‌های اعتبارسنجی فراهم ساخت.

^۱ Ensemble learning
^۲ Feature importance

^۳ Grid search
^۴ Fold cross-validation-10

۳-۳-۴- معیارهای ارزیابی عملکرد

عملکرد مدل در مجموعه آزمون با استفاده از شاخص‌های استاندارد طبقه‌بندی ارزیابی شد. این شاخص‌ها شامل دقت^۱، صحت پیش‌بینی مثبت^۲، نرخ بازیابی^۳ و میانگین هارمونیک آن‌ها^۴ است.

$$\begin{aligned} Accuracy &= \frac{TP + TN}{TP + TN + FP + FN} \\ Precision &= \frac{TP}{TP + FP} \\ Recall &= \frac{TP}{TP + FN} \\ F1 &= 2 \times \frac{Precision \times Recall}{Precision + Recall} \end{aligned} \quad (9)$$

علاوه بر این، منحنی مشخصه عملکرد گیرنده^۵ ترسیم و مقدار سطح زیر منحنی^۶ محاسبه شد تا توانایی تفکیک مدل در شناسایی شعب کارا و ناکارا به صورت جامع ارزیابی گردد.

۳-۳-۵- جایگاه مدل در چارچوب تصمیم‌یار

مدل *ML*، بخش پیش‌نگر سیستم تصمیم‌یار پیشنهادی را تشکیل می‌دهد. خروجی این مدل، احتمال ناکارایی شعب در دوره آتی است. شعبی که احتمال ناکارایی آن‌ها از آستانه تعیین شده مدیریتی فراتر رود، به عنوان شعب در معرض خطر شناسایی می‌شوند. این شعب در مرحله بعد وارد مدل *Inverse DEA* شده و مقادیر هدف خروجی‌ها برای انتقال به مرز کارایی تعیین می‌گردد.

بدین ترتیب، ترکیب *DEA*، *ML* و *Inverse DEA* چارچوبی یکپارچه فراهم می‌سازد که نه تنها عملکرد گذشته را ارزیابی می‌کند، بلکه امکان پیش‌بینی و مداخله اصلاحی مبتنی بر داده را نیز فراهم می‌آورد.

۴- داده‌ها و آمار توصیفی

در این پژوهش داده‌های عملیاتی ۴۰ شعبه یک بانک تجاری طی یک دوره چهارساله (۱۴۰۰ تا ۱۴۰۳) مورد استفاده قرار گرفت. هر ترکیب "شعبه-سال" به عنوان یک *DMU* مستقل در نظر گرفته شد؛ بنابراین، مجموعه داده شامل ۱۶۰ مشاهده است. داده‌ها از سامانه‌های اطلاعات مدیریت بانک استخراج و پس از حذف مقادیر گمشده و کنترل ناسازگاری‌های احتمالی، آماده تحلیل گردید.

انتخاب متغیرهای ورودی و خروجی بر اساس رویکرد واسطه‌گری^۷ در ارزیابی کارایی بانک‌ها به صورت جدول ۱ انجام گرفت؛ بدین معنا که شعب بانک به عنوان واسطه‌ای میان تجهیز منابع و تخصیص اعتبارات در نظر گرفته شدند.

جدول ۱- مشخصات ورودی‌ها و خروجی‌های سیستم بانکی.
Table 1- Specifications of inputs and outputs of the banking system.

عنوان	متغیرها	شرح	واحد
ورودی‌ها	X ₁	تعداد کارکنان شعبه	نفر
	X ₂	هزینه‌های عملیاتی سالانه	میلیون ریال
	X ₃	حجم سپرده‌های جذب شده	میلیون ریال
خروجی‌ها	Y ₁	حجم تسهیلات اعطایی	میلیون ریال
	Y ₂	درآمد کارمزدی	میلیون ریال
	Y ₃	تعداد حساب‌های فعال	عدد

¹ Accuracy² Precision³ Recall⁴ F1-score⁵ Receiver Operating Characteristic (ROC)⁶ Area Under the Curve (AUC)⁷ Intermediation approach

آمار توصیفی تمامی متغیرهای مورد استفاده در کل مشاهدات محاسبه شد و نتایج حاصل شامل شاخص‌هایی نظیر میانگین، میانه، حداقل و حداکثر متغیرها، در جدول ۲ ارائه شده است.

جدول ۲- آمار توصیفی متغیرها.

Table 2- Descriptive statistics of variables.

متغیر	میانگین	انحراف معیار	حداقل	حداکثر
تعداد کارکنان	13.4	3.2	6	22
هزینه عملیاتی	48.520	12.870	24.300	82.100
حجم سپرده‌ها	1,245.000	410.000	42.000	2,310.000
تسهیلات اعطایی	1,020.000	350.000	310.000	1,980.000
درآمد کارمزدی	74.300	21.500	28.400	132.600
تعداد حساب فعال	8.540	2.430	3.200	15.700

بررسی آمار توصیفی نشان می‌دهد پراکندگی قابل توجهی میان شعب وجود دارد که شرط تمایز پذیری در DEA را فراهم می‌سازد.

۵- نتایج تجربی

۵-۱- نتایج DEA

مدل BCC خروجی محور برای تمام ۱۶۰ واحد "شعبه-سال" اجرا شد که نتایج در جدول ۳ آورده شد.

جدول ۳- نتایج کارایی شعب در دوره ۱۴۰۳-۱۴۰۰.

Table 3- Branch efficiency results in the period 1400-1403.

سال	تعداد DMU	میانگین کارایی	انحراف معیار	حداقل کارایی	حداکثر کارایی	تعداد شعب کارا
1400	40	0.81	0.12	0.58	1.00	6
1401	40	0.83	0.11	0.61	1.00	7
1402	40	0.85	0.10	0.64	1.00	8
1403	40	0.87	0.09	0.68	1.00	8

نتایج جدول ۳ نشان می‌دهد میانگین کارایی شعب بانک طی دوره مورد بررسی روندی افزایشی داشته است. میانگین کارایی از ۰/۸۱ در سال ۱۴۰۰ به ۰/۸۷ در سال ۱۴۰۳ افزایش یافته که بیانگر بهبود تدریجی عملکرد عملیاتی شعب است. همچنین کاهش انحراف معیار کارایی در طول زمان نشان می‌دهد پراکندگی عملکرد میان شعب کاهش یافته و شبکه بانکی به سمت همگرایی عملکردی حرکت کرده است.

حداقل کارایی نیز از ۰/۵۸ به ۰/۶۸ افزایش یافته است که حاکی از آن است شعب بسیار ناکارا نیز در طول دوره مورد بررسی بهبود عملکرد داشته‌اند. تعداد شعب کارا نیز از ۶ شعبه در سال ۱۴۰۰ به ۸ شعبه در سال ۱۴۰۳ رسیده است.

این یافته‌ها نشان می‌دهد بخش قابل توجهی از ناکارایی شعب ناشی از ضعف مدیریتی و نحوه بهره‌برداری از منابع بوده و نه کمبود منابع؛ زیرا بدون افزایش محسوس ورودی‌ها، سطح خروجی‌ها بهبود یافته است.

از دیدگاه مدیریتی، نتایج نشان می‌دهد تمرکز بر بهبود فرایندهای عملیاتی، توسعه خدمات غیرحضور و مدیریت بهتر ارتباط با مشتریان می‌تواند تاثیر قابل توجهی بر کارایی شعب داشته باشد. همچنین مشاهده شد برخی شعب در تمامی سال‌ها روی مرز کارایی قرار داشته‌اند و به‌عنوان شعب مرجع در مدل DEA ظاهر شدند. این شعب می‌توانند به‌عنوان الگوی عملیاتی برای سایر شعب مورد استفاده قرار گیرند و انتقال تجربیات مدیریتی آن‌ها به سایر شعب می‌تواند یکی از راهکارهای کم‌هزینه بهبود عملکرد باشد.

۵-۲- نتایج مدل ML

مدل Random Forest با استفاده از داده‌های چهارساله آموزش داده شد (trained) و عملکرد آن بر روی مجموعه آزمون ارزیابی شد. نتایج ارزیابی با استفاده از داده‌های مطالعه حاضر در جدول ۴ ارائه شده است.

جدول ۴- نتایج مدل ML.
Table 4- ML model results.

مقدار	شاخص
0.88	Accuracy
0.86	Precision
0.90	Recall
0.88	F1-Score
0.92	AUC

مقدار بالای AUC نشان می‌دهد مدل توانایی مناسبی در تفکیک شعب کارا و ناکارا دارد.

تحلیل اهمیت متغیرها نشان داد سه عامل زیر بیشترین تاثیر را بر کارایی دارند:

۱. حجم سپرده‌ها.
۲. هزینه عملیاتی.
۳. تعداد حساب‌های فعال.

این یافته نشان می‌دهد مدیریت ترکیب منابع و کنترل هزینه‌ها نقش تعیین‌کننده‌ای در عملکرد شعب ایفا می‌کند.

۳-۵- نتایج مدل Inverse DEA

به منظور نشان دادن کاربرد مدیریتی مدل، شعبه شماره ۱۷ در سال ۱۴۰۳ بررسی شد. این شعبه دارای کارایی ۰/۷۲ بود و توسط مدل ML نیز به عنوان شعبه در معرض ناکارایی پیش‌بینی گردید.

اجرای مدل Inverse DEA نشان داد برای انتقال این شعبه به مرز کارایی، بدون افزایش ورودی‌ها، اهداف به شرح جدول ۵ باید محقق شود.

جدول ۵- نتایج مدل Inverse DEA.			
Table 5- Results of the Inverse DEA model.			
خروجی	مقدار فعلی	مقدار هدف	درصد تغییر
تسهیلات اعطایی	870.000	1,020.000	+17%
درآمد کارمزدی	61.200	70.800	+16%
حساب فعال	7.450	8.200	+10%

این نتایج نشان می‌دهد شعبه می‌تواند بدون افزایش تعداد کارکنان یا هزینه عملیاتی، صرفاً از طریق افزایش فعالیت اعتباری و جذب مشتریان فعال به مرز کارایی برسد.

۶- تحلیل و تفسیر نتایج

نتایج حاصل از اجرای مدل یکپارچه DEA، ML و Inverse DEA نشان می‌دهد کارایی شعب بانک صرفاً تابع میزان منابع در اختیار آن‌ها نیست، بلکه نحوه مدیریت و بهره‌برداری از منابع نقش تعیین‌کننده‌ای در عملکرد عملیاتی شعب دارد. اگرچه برخی شعب از نظر میزان سپرده و تعداد کارکنان در سطح مشابهی قرار داشتند، اما اختلاف قابل توجهی در کارایی مشاهده شد که بیانگر تفاوت در کیفیت مدیریت عملیاتی و ساختار ارایه خدمات است.

تحلیل اهمیت متغیرها در مدل ML نشان داد حجم سپرده‌ها، تعداد حساب‌های فعال و هزینه‌های عملیاتی مهم‌ترین عوامل موثر بر کارایی هستند. این نتیجه از دیدگاه اقتصادی قابل تفسیر است؛ زیرا افزایش سپرده‌ها ظرفیت اعطای تسهیلات را افزایش داده و از سوی دیگر افزایش تعداد حساب‌های فعال نشان‌دهنده پویایی ارتباط با مشتریان است. در مقابل، بالا بودن هزینه‌های عملیاتی بدون رشد متناسب در خدمات ارایه شده منجر به کاهش کارایی می‌شود.

همچنین نتایج *Inverse DEA* نشان داد بسیاری از شعب ناکارا برای رسیدن به مرز کارایی نیازمند افزایش منابع نیستند، بلکه باید ترکیب خدمات خود را اصلاح کنند. به‌طور مشخص، افزایش فعالیت‌های اعتباری و خدمات کارمزدی نسبت به افزایش تعداد کارکنان تاثیر بیشتری بر بهبود کارایی دارد. این موضوع بیانگر آن است که ناکارایی در شبکه بانکی بیشتر ناشی از ناکارایی مدیریتی است تا محدودیت منابع.

از منظر سیستم تصمیم‌یار، ترکیب سه روش مورد استفاده یک چرخه تصمیم‌گیری کامل ایجاد می‌کند. *DEA* وضعیت فعلی عملکرد را ارزیابی می‌کند، *ML* وضعیت آینده را پیش‌بینی می‌نماید و *Inverse DEA* اقدامات اصلاحی کمی و قابل اجرا ارائه می‌دهد؛ بنابراین، چارچوب پیشنهادی از یک ابزار سنجش عملکرد فراتر رفته و به یک سامانه پشتیبان تصمیم‌گیری مدیریتی تبدیل می‌شود که امکان پیش‌بینی و مداخله اصلاحی پیش‌دستانه را نیز فراهم می‌سازد.

۷- پیامدهای مدیریتی

نتایج حاصل از این پژوهش چند دلالت کاربردی برای مدیران شبکه بانکی دارد. نخست، ارزیابی کارایی شعب نباید صرفاً بر اساس شاخص‌های مالی منفرد انجام شود، بلکه لازم است مجموعه‌ای از شاخص‌های عملیاتی به‌طور هم‌زمان مورد بررسی قرار گیرد. دوم، شناسایی زود هنگام شعب در معرض افت عملکرد امکان مداخله پیشگیرانه مدیریتی را فراهم می‌سازد و می‌تواند از بروز زیان عملیاتی جلوگیری کند و سوم، اهداف تعیین شده توسط مدل *Inverse DEA* می‌تواند به‌عنوان مبنایی برای بودجه‌ریزی بر مبنای عملکرد شعب مورد استفاده قرار گیرد. به‌جای تعیین اهداف کلی، برای هر شعبه اهداف کمی متناسب با شرایط آن قابل تعریف است. همچنین بانک می‌تواند از شعب کارا به‌عنوان شعب الگو استفاده کرده و انتقال تجربیات مدیریتی میان شعب را تسهیل نماید.

۸- نتیجه‌گیری و پیشنهادهای آینده

در این پژوهش، یک چارچوب یکپارچه برای ارزیابی، پیش‌بینی و بهبود کارایی شعب بانک ارائه شد که شامل سه مولفه اصلی بود: *DEA*، مدل *ML* (*Random Forest*) و نتایج نشان داد میانگین کارایی شعب طی دوره ۱۴۰۰ تا ۱۴۰۳ روند افزایشی داشته و پراکندگی عملکرد میان شعب کاهش یافته است که نشان‌دهنده بهبود مستمر و همگرایی عملکردی شبکه بانکی است.

تحلیل اهمیت متغیرها در مدل *ML* نشان داد که حجم سپرده‌ها، تعداد حساب‌های فعال و هزینه‌های عملیاتی به‌عنوان مهم‌ترین عوامل موثر بر کارایی شعب شناخته شده‌اند. نتایج حاکی از آن است که ناکارایی شعب بیشتر با نحوه مدیریت منابع و بهره‌برداری از آن مرتبط است تا کمبود مطلق منابع. همچنین مدل *Inverse DEA* توانست اهداف عملیاتی کمی و قابل اجرا برای ارتقای شعب ناکارا ارائه دهد که این امر نشان‌دهنده کاربرد عملی و تصمیم‌یارانه چارچوب پیشنهادی است.

از نظر مدیریتی، این چارچوب به مدیران امکان می‌دهد تا با شناسایی زود هنگام شعب در معرض ناکارایی، تخصیص بهینه منابع و بهبود فرایندهای عملیاتی، بهره‌وری شبکه بانکی را افزایش دهند. همچنین شعب کارا می‌توانند به‌عنوان الگوی عملیاتی برای سایر شعب مورد استفاده قرار گیرند.

با توجه به یافته‌های این پژوهش، برخی مسیرهای توسعه و تحقیقات آینده قابل پیشنهاد هستند. این مسیرها شامل گسترش متغیرهای کیفی و شاخص‌های مشتری‌مداری، بهره‌گیری از الگوریتم‌های پیشرفته *ML*، تحلیل بلندمدت داده‌های پانل^۱ و اعمال چارچوب پیشنهادی در سایر موسسات مالی می‌باشد. اجرای این پیشنهادها می‌تواند قابلیت تعمیم و کاربردی بودن چارچوب را افزایش داده و مسیر تحقیقات بعدی در حوزه تحلیل کارایی و تصمیم‌گیری را هموار سازد.

^۱ Panel data

تشکر و قدردانی

نویسندگان از حسن همکاری مشارکت‌کنندگان در این پژوهش تشکر می‌کنند.

منابع مالی

پژوهشگران هیچ‌گونه حمایت یا کمک مالی جهت انجام و تکمیل این پژوهش دریافت نکرده‌اند.

تعارض با منافع

بنا بر اظهار نویسندگان این مقاله تعارض منافع ندارد.

منابع

- [1] Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research*, 2(6), 429–444. [https://doi.org/10.1016/0377-2217\(78\)90138-8](https://doi.org/10.1016/0377-2217(78)90138-8)
- [2] Banker, R. D., Charnes, A., & Cooper, W. W. (1984). Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management Science*, 30(9), 1078–1092. <https://doi.org/10.1287/mnsc.30.9.1078>
- [3] Emrouznejad, A., & Yang, G. liang. (2018). A survey and analysis of the first 40 years of scholarly literature in DEA: 1978–2016. *Socio-Economic Planning Sciences*, 61, 4–8. <https://doi.org/10.1016/j.seps.2017.01.008>
- [4] Ullah, S., Majeed, A., & Popp, J. (2023). Determinants of bank's efficiency in an emerging economy: A data envelopment analysis approach. *Plos One*, 18(3), e0281663. <https://doi.org/10.1371/journal.pone.0281663>
- [5] Zhong, K., Li, C., & Wang, Q. (2021). Evaluation of bank innovation efficiency with data envelopment analysis: From the perspective of uncovering the black box between input and output. *Mathematics*, 9(24), 3318. <https://doi.org/10.3390/math9243318>
- [6] Altındağ, İ. (2024). Evaluating the efficiency of Turkish banking sector using data envelopment analysis and Malmquist productivity index: Covid-19 pandemic and after. *Pakistan Journal of Statistics and Operation Research*, 20(4), 719–731. <http://dx.doi.org/10.18187/pjsor.v20i4.4486>
- [7] Dere, İ., Turanlı, M., Alp, S., & Erdoğan, M. F. (2025). Analysis impact of financial ratios on bank success using machine learning classification algorithms: The case of Turkey. *Journal of Soft Computing and Decision Analytics*, 3(1), 50–71. <https://doi.org/10.31181/jsda31202553>
- [8] Hoque, A., Islam, I., Chowdhury, S. A., Mahmud, I., & Hossain, A. J. (2025). AI and machine learning in banking: Driving efficiency and innovation. *Well Testing Journal*, 34(S3), 80–101. <https://welltestingjournal.com/index.php/WT/article/view/176>
- [9] Lim, D. J. (2020). Inverse data envelopment analysis for operational planning: The impact of oil price shocks on the production frontier. *Expert Systems with Applications*, 161, 113726. <https://doi.org/10.1016/j.eswa.2020.113726>
- [10] Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32. <https://doi.org/10.1023/A:1010933404324>
- [11] Amirteimoori, A., & Allahviranloo, T. (2025). Performance and managerial ability analysis in banking sector: A fuzzy data envelopment analysis approach. *Soft Computing*, 29(7), 3289–3305. <https://doi.org/10.1007/s00500-025-10556-5>
- [12] Gherțescu, C., Manta, A. G., Bădîrcea, R. M., & Manta, L. F. (2024). How does the digitalization strategy affect bank efficiency in Industry 4.0? A bibliometric analysis. *Systems*, 12(11), 492. <https://doi.org/10.3390/systems12110492>