

## انتخاب شاخص‌های ارزیابی در تحلیل پوششی داده‌ها با استفاده از آنالیز تشخیصی

محمد رحیم رمضانیان\* - کیخسرو پاکیده\*\* - اکرم اویسی عمران\*\*\*

(تاریخ دریافت: ۹۲/۴/۳۱ - تاریخ پذیرش: ۹۳/۱/۱۰)

### چکیده

انتخاب متغیرهای ورودی و خروجی در تحلیل پوششی داده‌ها از اهمیت به‌سزایی برخوردار است و نمرات کارایی واحدها به شدت از مجموعه متغیرهای انتخابی تأثیر می‌پذیرند. در این پژوهش، روشی برای انتخاب متغیرهای ورودی و خروجی، براساس آنالیز تشخیصی ارائه شده است. در این روش اهمیت نسبی متغیرها در تعیین عضویت واحدها در دو گروه متغیر وابسته (کارا و ناکارا) یا همان ضرایب تشخیصی استاندارد به‌عنوان معیار انتخاب متغیرهای اصلی مدل معرفی می‌شوند.

واژگان کلیدی: تحلیل پوششی داده‌ها، آنالیز تشخیصی، ضرایب استاندارد تشخیصی، تابع تشخیصی.

---

\* دانشیار گروه مدیریت، دانشکده ادبیات و علوم انسانی، دانشگاه گیلان

\*\* استادیار گروه مدیریت، دانشکده ادبیات و علوم انسانی، دانشگاه گیلان

\*\*\* کارشناس ارشد مدیریت صنعتی، گروه مدیریت، دانشکده ادبیات و علوم انسانی، دانشگاه گیلان (نویسنده مسئول)

akramoveysi@gmail.com

مقدمه

فارل<sup>۱</sup> (۱۹۵۷) برای اولین بار تخمین کارایی به روش ناپارامتریک را مطرح کرد (Farrell 1957). او به جای تخمین تابع تولید، مرزی را برای واحدهای تصمیم گیرنده<sup>۲</sup> در نظر گرفت که این مرز را ملاک سنجش کارایی قرار داد. اما تحلیل پوششی داده‌ها<sup>۳</sup> با تز دکتری «ادوارد رودز»<sup>۴</sup> تحت راهنمایی «کوپر و چارنز»<sup>۵</sup> شروع شد که در آن پیشرفت تحصیلی دانش آموزان مدارس آمریکا را در سال ۱۹۷۸ مورد ارزیابی قرار داده بودند و این تحقیق در نهایت به معرفی مدل CCR<sup>۶</sup> منتج گردید (Charnes and et all, 1978). این مدل با فرض بازده به مقیاس ثابت<sup>۷</sup> (CRS) و با نگرش به ورودی‌ها طراحی شده بود. اما فرض بازده به مقیاس ثابت زمانی مناسب است که همه بنگاه‌ها در مقیاس بهینه عمل کنند. بنکر، چارنز و کوپر<sup>۸</sup> (۱۹۸۴) مدل CCR را بسط داده و فرض بازده به مقیاس متغیر<sup>۹</sup> (VRS) را به آن افزودند که منجر به ارائه مدل BCC<sup>۱۰</sup> شد (Charnes and et all, 1984).

Min  $\theta$

$$\begin{aligned} \text{s.t. : } & \sum_{j=1}^n \lambda_j x_{ij} \leq \theta x_{i0} \quad , i = 1, 2, \dots, m \\ & \sum_{j=1}^n \lambda_j y_{rj} \geq y_{r0} \quad , r = 1, 2, \dots, s \\ & \sum_{j=1}^n \lambda_j = 1 \\ & \lambda_j \geq 0 \quad , j = 1, 2, \dots, n \\ & \theta : \text{ free} \end{aligned} \quad (1)$$

1. Farrell
2. Decision Making Units (DMUs)
3. Data Envelopment Analysis (DEA)
4. Edward Rhodes
5. Cooper And Charnes
6. Charnes And Cooper And Rhodes (CCR)
7. Constant Return To Scale (CRS)
8. Banker, Charnes And Cooper
9. Variable Return to Scale(VRS)
10. Banker, Charnes and Cooper (BCC)

انتخاب متغیرهای ورودی و خروجی، در ارزیابی کارایی یک واحد از اهمیت بالایی برخوردار است. در عمل تعداد زیادی از متغیرها را می‌توان به‌عنوان عوامل مؤثر بر کارایی در نظر گرفت. رویکردهای مختلفی برای تعیین ورودی‌ها و خروجی‌های بانک‌ها و مؤسسات مالی وجود دارد. اما عمده‌ترین آن‌ها دو رویکرد تولیدی و رویکرد واسطه‌ای هستند. در رویکرد واسطه‌ای، فرض می‌شود بانک‌ها از نیروی کار، سرمایه و سپرده‌ها برای تولید تسهیلات مالی استفاده نمی‌کنند.

اما در رویکرد تولیدی علاوه بر وام، سپرده‌ها نیز به‌عنوان خروجی سیستم در نظر گرفته می‌شود، به‌طوری‌که ورودی‌ها فقط شامل نیروی کار و سرمایه است (Berger and Humphrey 1997). صرف نظر از تشخیص ماهیت یک متغیر از حیث ورودی یا خروجی، معمولاً تعداد متغیرهای مؤثر بر عملکرد یک سیستم بسیار زیاد هستند. در یک مدل DEA با هر منبع به کار گرفته شده برای یک واحد باید به‌عنوان یک متغیر ورودی رفتار شود. متغیرهای خروجی هم در نتیجه فعالیت بنگاه به‌منظور تبدیل ورودی‌ها به خروجی‌ها به کار می‌روند. همچنین در برخی مواقع لازم است متغیرهای محیطی که بر فعالیت واحد تأثیر گذارند، نیز در فهرست متغیرها وارد شوند (Klimberg and Puddicombe, 1995).

آن دسته از متغیرهای محیطی که منابعی را به بنگاه می‌افزایند به‌عنوان ورودی و آن دسته از متغیرهایی که نیازمند منابع بنگاه هستند به‌عنوان خروجی در نظر گرفته می‌شوند (Pinto and et all, 2012).

استفاده از تعداد زیاد متغیرهای ورودی و خروجی دقت تحلیل را کاهش می‌دهد (Jenkins and Anderson, 2003). به‌علاوه تعداد زیاد متغیرها باعث کاهش تفاوت بین نمرات کارایی واحدها شده و به این ترتیب تعداد بیشتری از واحدها کارا به‌نظر خواهند رسید (Golany and roll, 1989). در واقع تعداد بیشتر متغیرها در DEA علاوه بر افزایش نمرات کارایی، تعداد واحدهای کارا را نیز افزایش داده و به کاهش قدرت تشخیص مدل می‌انجامد (Nunamaker, 1985). برای شناسایی مرتبط‌ترین متغیرها در مدل‌های DEA، تاکنون رویکردهای متعددی توسعه یافته‌اند. یکی از این رویکردها شامل غربال‌گری فهرست متغیرها براساس نظر متخصصان است که از آن جمله می‌توان به روش دلفی یا روش تحلیل سلسله‌مراتبی اشاره نمود (Golany and roll, 1989).

رویکرد دیگری که برای کاهش متغیرها به کار گرفته می‌شود، روش تحلیل رگرسیون و همبستگی بین متغیرها است. در این رویکرد همبستگی بالا بین متغیرهای موجود در مدل به معنی

زاید بودن متغیرها تلقی می‌شود. به عبارتی در این روش هدف، انتخاب مجموعه متغیرهایی است که کم‌ترین همبستگی را با یکدیگر داشته باشند (Salinas and et all, 1996). اما همبستگی آماری بین متغیرها نمی‌تواند معیار مناسبی برای انتخاب متغیرهای ورودی و خروجی باشد چراکه زمانی که یک متغیر در تحلیل همبستگی متغیری زاید است، لزوماً در مدل DEA نیز محدودیت زاید محسوب نمی‌شود. به بیان دیگر وجود همبستگی بالا بین متغیرها به این معنا نیست که اضافه شدن متغیر در تحلیل بر نمرات کارایی تأثیر نخواهد داشت. بنابراین کاربرد تحلیل همبستگی برای کاهش متغیرها اگرچه قبلاً مورد استقبال قرار گرفته بود، اما بعدها مورد انتقاد قرار گرفت در مدل DEA غیرمنطقی خواهد بود (Nunamaker, 1985).

در مقابل این روش‌ها، در رویکرد دیگری، اثر متغیرها بر روی تغییر نمرات کارایی که با بررسی اثر حذف یا کاربرد آن سنجیده می‌شود، مبنا قرار می‌گیرد. با استفاده از این رویکرد، به منظور بررسی اثر حذف یا کاربرد متغیر بر نمرات کارایی از آزمون آماری که توسط بنکر معرفی شده است، استفاده می‌شود (Banker, 1993)، (Banker, 1996). هنگامی که تغییرات نمرات کارایی از نظر آماری معنی دار باشد، آنگاه می‌توان در مورد مؤثر بودن متغیر بر نمرات کارایی مطمئن بود و از حذف آن از مدل خودداری کرد (Edirisinghe and Zhang, 2010)، (Kittelson, 1996)، (Yan and et all, 2012).

در این پژوهش به منظور انتخاب مؤثرترین متغیرهای ورودی و خروجی از رویکرد جدیدی براساس روش آنالیز تشخیصی<sup>۱</sup> استفاده شده است.

### مبانی نظری تحقیق

نظریه ابتدایی آنالیز تشخیصی به دهه ۱۹۳۰ و آماردان انگلیسی کارل پیرسون<sup>۲</sup> و همکاران، در زمینه فواصل گروه‌ها یا ضرایب تشابه نژادی برمی‌گردد. اما این روش به طور خاص توسط فیشر (۱۹۳۶) ابداع شد و بر پایه جبر ماتریسی مورد استفاده در رگرسیون خطی چندمتغیره برای حل معادلات خطی توسعه یافت (منصورفر، ۱۳۸۵). آنالیز تشخیصی برای طبقه‌بندی پاسخگویان براساس مقادیر (کدهای) یک متغیر وابسته اسمی دو یا چندوجهی به کار می‌رود. در مواردی که متغیر وابسته اسمی و متغیرهای مستقل کمی باشند، برای پیش‌بینی تغییرات متغیر وابسته از روی

1. Discriminant Analysis (DA)

2. Charle Pearson

متغیرهای مستقل، می‌توان از آنالیز تشخیصی استفاده کرد (pardo and et all, 2007). با فرض این که متغیر وابسته از نوع اسمی دوجهی باشد و بتوان آن را با کدهای صفر و یک مشخص کرد، این متغیر مفهوم عضویت در گروه‌های دوگانه می‌دهد. در این شرایط هدف آنالیز تشخیصی پیدا کردن ترکیبی خطی از متغیرهای مستقل کمی است به گونه‌ای که این ترکیب منجر به بهترین پیش‌بینی ممکن از نمرات متغیر وابسته که با کدهای دوگانه مشخص شده گردد که این ترکیب خطی، تابع تشخیصی نام دارد (pinto and et all, 2012).

### مفروضات آنالیز تشخیصی

- ۱- متغیر وابسته یک متغیر اسمی دوجهی یا چندوجهی واقعی است. توصیه می‌شود که هیچ‌گاه برای انجام آنالیز تشخیصی یک متغیر کمی فاصله‌ای/نسبی به صورت ساختگی به یک متغیر دوجهی تبدیل نشود.
- ۲- تمامی واحدهای تصمیم‌گیرنده باید مستقل از یکدیگر باشند. بنابراین نمی‌توان از داده‌های مرتبط (مانند: داده‌های قبل و بعد، داده‌های پانل و داده‌های جفتی) استفاده نمود.
- ۳- تمامی پاسخگویان باید به گروه‌هایی که براساس متغیر وابسته تشکیل می‌شوند، تعلق گیرند. گروه‌ها باید به‌طور متقابل مانع باشند و هر پاسخگو به یک گروه اختصاص یابد.
- ۴- حداقل باید دو پاسخگو برای هر گروه وجود داشته باشد.
- ۵- حداکثر تعداد متغیر مستقلی که می‌توانیم استفاده کنیم برابر با تعداد نمونه منهای ۲ ( $n-2$ ) است.
- ۶- انحراف استاندارد هیچ‌یک از متغیرهای مستقل در گروه‌هایی که تشکیل می‌شوند برابر صفر نمی‌باشد.
- ۷- واریانس هر کدام از متغیرهای مستقل در بین گروه‌هایی که براساس متغیر وابسته شکل می‌گیرند، باید برابر باشد. عدم برابری واریانس‌ها می‌تواند نشان‌دهنده وجود داده‌های پرت در یک یا چند گروه باشد.
- ۸- همبستگی بین هر دو متغیر مستقل در هر گروه باید با همبستگی بین هر دو متغیر مستقل در گروه دیگر برابر باشد. بنابراین هر گروهی یک ماتریس همبستگی مشابه و برابر دارد.
- ۹- هم‌خطی کامل بین متغیرهای مستقل وجود نداشته باشد. یعنی یکی از متغیرهای مستقل همبستگی بالایی با متغیرهای دیگر داشته باشد و یا این که یکی از متغیرهای مستقل تابع متغیرهای

دیگر باشد، مجاز نیست. در آن صورت مقدار تولرانس برای آن متغیر به عدد صفر نزدیک شده و ماتریس به دست آمده یک تابع تشخیص منحصر به فرد نخواهد بود. هر چه میزان همبستگی متغیرهای مستقل بیشتر باشد، تابع تشخیصی از توان کمی برخوردار است (Tabachnick and Fidell, 2011)، (Warner, 2013).

### مفاهیم کلیدی آنالیز تشخیصی

متغیرهای تمیزدهنده ۱: همان متغیرهای مستقل هستند که متغیرهای پیشین نیز به آن‌ها اطلاق می‌شود.

متغیر ملاک یا معیار ۲: همان متغیر وابسته است که به عنوان متغیر گروه‌بندی شده در SPSS معروف است. هدف اصلی طبقه‌بندی در آنالیز تشخیصی نیز همین متغیر است.

تابع تشخیصی ۳: تابع تشخیصی که به ریشه کانونی ۴ معروف است، ترکیب خطی وزنی بهینه از نمرات متغیرهای مستقل است. تابع تشخیصی ممکن است از نمرات خام متغیرهای مستقل  $X$  و یا نمرات استاندارد  $Z$  محاسبه شود. تابع تشخیصی مشابه معادله رگرسیونی است. یک نمره تابع تشخیصی برای هر واحد از روی مجموعه‌ی متغیرهای مستقل که به وسیله یک ضریب وزن داده شده است، محاسبه می‌گردد. جایی که دو گروه وجود داشته باشد، فقط یک تابع تشخیصی وجود خواهد داشت و در نتیجه یک مجموعه از ضرایب تابع تشخیصی برای آن وجود دارد. رایج‌ترین شکل نمایش تابع تشخیصی به صورت زیر است:

$$D_i = d_{i1}Z_1 + d_{i2}Z_2 + \dots + c_{ip}X_p \quad (۲)$$

مقدار این تابع از حاصل ضرب نمرات استاندارد متغیرهای مستقل یعنی  $Z$  در ضریب تابع تشخیصی استاندارد  $d_i$  محاسبه می‌شود.

ضرایب تابع تشخیصی بر اساس مفهوم همبستگی کانونی محاسبه می‌شوند. مفهوم همبستگی کانونی در واقع به مفهوم پیدا کردن ترکیب خطی از متغیرهای مستقل و ترکیب خطی از متغیرهای وابسته است که حداکثر همبستگی را داشته باشند. تحلیل تشخیصی نیز اساساً یک مسئله همبستگی

1. Discriminating variables (DVS)
2. Criterion variables (CV)
3. Discriminant function (DF)
4. Canonical Root

کانونی است که عضویت در گروه‌ها در یک سمت معادله و متغیرهای مستقل در طرف دیگر آن قرار دارند. محاسبه همبستگی کانونی و ترکیب‌های خطی از متغیرهای مستقل و وابسته که بیشترین همبستگی را ایجاد کنند از طریق محاسبه بردار ویژه و مقدار ویژه ماتریس همبستگی ممکن می‌شود. در تحلیل تشخیص وقتی تنها دو طبقه در مسئله مطرح است، در واقع به پیدا کردن ترکیب خطی از متغیرهای مستقل که بیشترین همبستگی با متغیر وابسته با کد دوگانه تبدیل می‌شود. به بیان دیگر ضرایب در سمت راست معادله (۲) به گونه‌ای انتخاب می‌شوند که مقدار تابع بیشترین همبستگی، با متغیر وابسته که بیانگر تعلق به دو گروه است داشته باشد. به این ترتیب برای هر واحد مورد بررسی یک نمره تشخیصی  $D_i$  قابل محاسبه است. به دلیل این که میانگین هر متغیر زمانی که استاندارد شده باشد، صفر و انحراف استاندارد آن برابر با ۱ است، میانگین نمره تشخیصی برای تمام واحدها برابر با صفر است. بنابراین واحدهایی که نمره تشخیصی آن منفی باشد در یک گروه و واحدهایی که نمره تشخیصی آن‌ها مثبت باشد در گروه دیگر طبقه‌بندی می‌شوند (Tabachnick and Fidell, 2011).

نکته: در مواقعی که متغیر وابسته دوجویی باشد، تعداد تابع تشخیصی همواره برابر با یک است. اما زمانی که متغیر وابسته چندوجهی باشد، در آن صورت تعداد تابع با حداقل یکی از مقادیر زیر برابر است (Warner, 2013).

۱- تعداد طبقات متغیر وابسته منهای یک ( $k-1$ )

۲- تعداد متغیرهای مستقل تمیزدهنده

کاربرد ضرایب تشخیصی استاندارد: این ضرایب، که به ضرایب استاندارد شده تابع تشخیصی کانونی نیز معروف‌اند، برای مقایسه اهمیت نسبی متغیرهای مستقل به کار می‌روند (Tabachnick and Fidell, 2011)، (Warner, 2013).

### ارزیابی مدل یا تابع آنالیز تشخیصی

برای آزمون کارایی تابع تشخیصی در ایجاد تفاوت‌های معنی‌دار بین گروه‌ها، از آماره‌ای به نام لاندای ویلکز<sup>۱</sup> استفاده می‌شود که معنی‌داری آن از روی مقدار آماره کای اسکوئر به دست می‌آید (Warner, 2013).

### انتخاب متغیرهای ورودی و خروجی

پس از بررسی مفروضات آنالیز تشخیصی، با توجه به ضرایب تشخیصی استاندارد که اهمیت نسبی متغیرهای موجود در مدل را نشان می‌دهند؛ متغیرهای اصلی مدل انتخاب می‌شوند. هدف این روش انتخاب یک متغیر ورودی و خروجی است که دارای بیشترین میزان تأثیر در گروه‌بندی واحدها در دو گروه متغیر وابسته Y هستند.

### بحث و یافته‌های تحقیق

در این پژوهش ابتدا با استفاده از آنالیز تشخیصی متغیرهای مدل کاهش یافته و سپس کارایی ۳۶ شعبه درجه ۳ اداره سرپرستی منطقه جنوب بانک سپه سنجیده شد. با استفاده از ادبیات موضوع و نظر کارشناسان و مدیریت اداره سرپرستی منطقه جنوب بانک سپه، فهرست اولیه‌ای از متغیرهای پژوهش انتخاب گردید که ۸ متغیر اولیه عبارتند از:

متغیرهای ورودی: سود پرداختنی (سود پرداختی به انواع سپرده‌های گران‌قیمت)، هزینه کل (هزینه اداری و پرسنلی) و میزان سرمایه (اعم از منقول و غیرمنقول).

متغیرهای خروجی: میزان تسهیلات پرداختی، درآمد کل، میزان سپرده کل، نسبت تسهیلات به مطالبات و تعداد اسناد.

### انتخاب متغیرهای ورودی و خروجی

#### تعیین حجم نمونه

برای انجام آنالیز تشخیصی به‌ازای هر متغیر مستقل از بین کل شعب بانک سپه شهر تهران، ۱۸ واحد تصمیم‌گیرنده به‌عنوان نمونه انتخاب شد. به‌عبارت دیگر به‌ازای ۸ متغیر موجود در مدل، ۱۴۴ شعبه از شعب موجود در شهر تهران به‌عنوان نمونه انتخاب و نمرات کارایی این شعب با استفاده از فرم BCC ورودی‌گرا و با استفاده از متغیرهای اولیه محاسبه شدند (جهانشاهلو و همکاران، ۱۳۹۱)، (Tone and et all, 2010). این نمرات کارایی تعلق هر واحد به یکی از دو گروه متغیر وابسته را مشخص می‌کنند.

### بررسی مفروضات آنالیز تشخیصی

۱- متغیر کارایی که یک متغیر اسمی دوجهبی واقعی است که در دو سطح کارا و ناکارا طبقه‌بندی می‌شود.



efficiency = 1  $\Rightarrow$  coded by 1 , efficiency  $\neq$  1  $\Rightarrow$  coded by 2

۲- تمامی واحدهای تصمیم‌گیرنده مستقل از یکدیگر بوده و از بین ۱۷۸ شعبه اداره سرپرستی منطقه جنوب شهر تهران ۱۴۴ شعبه برای نمونه به روش نمونه‌گیری تصادفی انتخاب می‌شوند.

۳- نمرات کارایی واحدهای تصمیم‌گیرنده با استفاده از مدل BCC ورودی‌گرا محاسبه گردید و واحدها به گروه‌های کارا و ناکارا تفکیک شدند. گروه‌ها به‌طور متقابل مانع بوده و هر واحد به یک گروه اختصاص می‌یابد. در جدول ۱، احتمالات اولیه برای عضویت واحدها در دو گروه کارا و ناکارا را نشان می‌دهد که برآوردی از احتمال تعلق یک واحد به یک گروه خاص است. در حالتی که هیچ‌گونه اطلاعی در مورد عضویت واحدها نداریم، فرض بر این است که یک واحد با احتمال برابر ۰,۵ می‌تواند به هر گروهی تعلق داشته باشد.

جدول ۱- احتمالات پیشین برای هر گروه

کارایی	قبلی
۱	۰,۵
۲	۰,۵
نهایی	۱

۴- حداقل باید دو واحد برای هر گروه وجود داشته باشد. در این تحقیق ۸۱ شعبه از شعب نمونه کارا و ۶۳ شعبه ناکارا است. بنابراین این فرض نیز در مورد این نمونه رعایت شده است.

۵- حداکثر تعداد متغیر مستقلی که می‌توان استفاده نمود، برابر با تعداد نمونه منهای ۲ (n-2) است. در این تحقیق تعداد ۸ متغیر مستقل اولیه با توجه به ادبیات موضوع و نظر کارشناسان اداره حسابداری و آموزش بانک سپه به‌عنوان ورودی و خروجی انتخاب شدند.

۶- انحراف استاندارد هیچ‌یک از متغیرهای ورودی و خروجی در گروه‌های کارا و ناکارا برابر صفر نیست. در جدول ۲، مقدار آماره F نشان‌دهنده برابری میزان میانگین‌های دو گروه متغیر وابسته کارایی می‌باشد. آماره F دارای دو درجه آزادی است که df1 درجه آزادی صورت و df2 درجه آزادی مخرج است. در سطح معنی‌داری کوچک‌تر از ۰,۱، آماره‌های کوچک‌تر از ۰,۱ بیانگر وجود تفاوت معنی‌دار بین دو گروه است. همان‌طور که در این جدول مشاهده می‌شود، در بین میانگین‌های دو گروه برای ۵ متغیر سود پرداختی، هزینه کل، درآمد کل، نسبت تسهیلات به مطالبات و تعداد اسناد تفاوت معنی‌داری وجود ندارد.

جدول ۲- انحراف معیار متغیرهای مستقل

Sig	Std. deviation	df2	df1	F	
۰,۲۵۳	۶۴۶	۱۴۲	۱	۰,۸۱۱	سود پرداختنی
۰,۱۰۸	۱۶۹	۱۴۲	۱	۹,۲۴۶	هزینه کل
۰,۰۰۰	۲۲۵۴۶	۱۴۲	۱	۱۶,۰۴۵	میزان سرمایه
۰,۰۰۲	۱۲۶۴۱	۱۴۲	۱	۱۰,۱۲۵	میزان تسهیلات
۰,۷۵۵	۶۴۲	۱۴۲	۱	۰,۰۹۷	درآمد کل
۰,۰۱۷	۲۳۱۹۰	۱۴۲	۱	۵,۸۷۴	سپرده کل
۰,۸۲۲	۵۸	۱۴۲	۱	۰,۰۵۱	نسبت تسهیلات به مطالبات
۰,۱۳۷	۲۷۷۴	۱۴۲	۱	۰,۴۳۴	تعداد اسناد

۷- در داخل هر گروهی که براساس متغیر وابسته شکل می‌گیرد، واریانس هر یک از متغیرهای مستقل باید در دو گروه مشابه باشد. آزمونی که بدین منظور به کار می‌رود، آزمون لوین ۱ است. چنانچه سطح معنی‌داری در آزمون لوین از مقدار ۰,۰۵ بزرگ‌تر باشد، در آن صورت فرض برابری واریانس‌ها پذیرفته‌است. اما چنانچه سطح معنی‌داری در آزمون کوچک‌تر از این مقدار باشد، فرض برابری واریانس‌ها پذیرفته نیست.

جدول ۳- آزمون برابری واریانس‌ها

	آزمون لوین	
	F	sig
فرض برابری واریانس‌ها	۱۸,۳۰۷	۰,۱۲۱
فرض نابرابری واریانس‌ها	-	-

۸- آزمون برابری کوواریانس‌های دو گروه، آزمون ام. باکس ۲ است که به آزمون فرضیه‌های زیر می‌پردازد:

1. Levene
2. Box's M Test

کوواریانس دو گروه برابر هستند:  $H_0$

کوواریانس دو گروه برابر نیستند:  $H_1$

معنی‌داری این آزمون براساس مقدار تبدیل شده  $F$  است. در تفسیر نتیجه این آزمون باید گفت که اگر سطح معنی‌داری کوچک‌تر از  $0,1$  باشد، فرض برابری کوواریانس‌ها رد می‌شود و ماتریس‌ها با هم تفاوت دارند. چنانچه سطح معنی‌داری از مقدار  $0,1$  بزرگ‌تر باشد نیز دال بر برابری ماتریس‌های کوواریانس است. همان‌طور که در جدول ۵ مشاهده می‌شود؛ میزان معنی‌داری ( $0,293$ ) از ( $0,1$ ) بزرگ‌تر است، بنابراین فرض صفر پذیرفته شده و ماتریس کوواریانس دو گروه با هم برابر هستند.

جدول ۴- آزمون Box's M

Box's M		۱۳۱,۲۷۷
F	Approx.	۱,۳۸۰
	df1	۳۶
	df2	۲۱۴۱۹,۷
	Sig.	۰,۲۹۳

۹- جدول بعدی لگاریتم دترمینان نام دارد. این جدول نشان می‌دهد که ماتریس کوواریانس کدام یک از گروه‌های کارا و ناکارا بیش‌ترین تفاوت را با ماتریس کوواریانس داخل گروهی کل دارد. بنابراین هر گروهی یک لگاریتم دترمینان دارد که مقدار آن، نتیجه مقادیر ویژه ماتریس کوواریانس درون گروهی همان گروه است که در این جدول تفاوت قابل توجهی با یکدیگر ندارند و احتمال برابری کوواریانس دو گروه را افزایش می‌دهد. ستون رتبه نشان دهنده حداقل تعداد متغیرهای مستقلی است که با یکدیگر رابطه هم‌خطی ندارند. با توجه به وجود ۸ متغیر مستقل در این پژوهش، ستون رتبه نشان می‌دهد که تمامی متغیرها با یکدیگر رابطه هم‌خطی ندارند.

جدول ۵- رتبه و لگاریتم طبیعی گروه‌ها

کارایی	رتبه	لگاریتم دترمینان
گروه ۱	۸	۱۲۹,۷۵۷
گروه ۲	۸	۱۳۲,۳۳۰
داخل گروهی کل	۸	۱۳۱,۱۰۱

### ارزیابی مدل آنالیز تشخیصی

لانداى ويلكز عبارت است از نسبتى از واريانس و کوواريانس كل در نمرات متغيرهاى مستقل، كه براساس تفاوت بين گروهها قابل تبیین نمی باشد. هر چه ميزان لاندا كوچك تر باشد، نشان دهنده وجود تفاوت بين گروهها است. به عبارتى لانداى ويلكز (۰,۴۶۸-۱) از واريانس متغير وابسته كارایی را تبیین کرده است. بنابراین در سطح معنی داری صفر میانگین گروهها با هم متفاوتند.

جدول ۶- لانداى ويلكز

Test of Function(s)	Wilks' Lambda	Chi-square	df	Sig.
۱	۰,۴۶۸	۶۰,۰۱۷	۸	۰,۰۵۸

### تابع تشخیصی

جدول ۷ با عنوان «ضرایب تابع تشخیص کانونی»، ضرایب تشخیصی استاندارد برای هر متغیر مستقل است و سهم تفکیکی هر متغیر مستقل در تابع تشخیصی را ضمن کنترل اثر سایر متغیرها نشان می دهد.

جدول ۷- ضرایب تابع تشخیص کانونی

متغیرها	تابع
	۱
سود پرداختنی	۰,۹۷۴
هزینه های کل	۰,۷۶۷
میزان سرمایه	۰,۹۶۶
میزان تسهیلات	-۰,۴۴۹
درآمد کل	-۰,۷۸۱
میزان سپرده ها	-۰,۰۸۷
نسبت تسهیلات به مطالبات	-۰,۴۰۴
تعداد اسناد	-۰,۶۷۲
Constant	-۶,۷۲۵

### طبقه بندی نتایج تحلیل تشخیصی

جدول ۸ به عنوان مهم ترین جدول طبقه بندی واحدها در دو گروه کارا و ناکارا، میزان موفقیت طبقه بندی انجام شده را نشان می دهد. در این جدول که جدول آشفته گی نام دارد،

ردیف‌ها طبقات مشاهده‌شده متغیر وابسته یعنی کارا و ناکارابودن واحد در ارزیابی تحلیل پوششی داده‌ها و ستون‌ها نیز طبقات پیش‌بینی شده آن یعنی مثبت یا منفی بودن مقدار تابع تشخیص هستند. با توجه به این جدول از یک گروه ۸۱ واحدی از واحدهای کارا، ۶۰ واحد (معادل ۱,۷۴ درصد) از واحدها به درستی طبقه‌بندی شده‌اند و ۲۱ واحد (معادل ۹,۲۵ درصد) هم به اشتباه در گروه واحدهای کارا قرار گرفته‌اند. از بین ۶۳ واحد ناکارا، ۴۷ واحد (معادل ۶,۷۴ درصد) به درستی طبقه‌بندی شده‌اند و ۱۶ واحد (معادل ۴,۲۵ درصد) هم به اشتباه در گروه واحدهای ناکارا قرار گرفته‌اند. بنابراین ملاحظه می‌شود که دقت طبقه‌بندی در دو گروه واحدهای ناکارا و کارا تقریباً برابر است. دقت کل پیش‌بینی نسبت ضربه ۱ نامیده می‌شود و هر چه مقدار آن بیش‌تر باشد، دقت طبقه‌بندی و تفکیک گروه‌ها از یکدیگر بیش‌تر است. در مجموع براساس این نسبت در دو گروه ۴,۸۶۲ درصد به درستی در دو گروه کارا و ناکارا طبقه‌بندی شده‌اند و ۶,۱۳۳ درصد باقی‌مانده نیز به اشتباه طبقه‌بندی شده‌اند.

جدول ۸- طبقه‌بندی نتایج (آشفتگی)

گروه‌بندی	گروه	عضویت گروهی پیش‌بینی شده		نهایی
		۴۱	۵۲	
اصلی	۱	۶۰	۲۱	۸۱
	تعداد			
	۲	۱۶	۴۷	۶۳
	%			
	۱	۷۴,۱	۲۵,۹	۱۰۰
	۲	۲۵,۴	۷۴,۶	۱۰۰
معتبر متقاطع	۱	۶۰	۲۱	۸۱
	تعداد			
	۲	۱۶	۴۷	۶۳
	%			
	۱	۷۴,۱	۲۵,۹	۱۰۰
	۲	۲۵,۴	۷۴,۶	۱۰۰

1. Hit Ratio
2.  $[(60 \times 47) / 144] / 100 = 86.4\%$
3. 100- 86.4
4. Efficient Units
5. Inefficient Units

### ارزیابی کارایی واحدها

در نهایت با توجه به رعایت مفروضات و تحلیل دقت آنالیز تشخیصی، همچنین با در نظر گرفتن ضرایب تابع تشخیص کانونی (جدول ۷)، متغیرهای ورودی و خروجی مطابق با جدول ۹ انتخاب شدند. مشاهده می‌شود که دو متغیری که دارای بیشترین ضریب تشخیصی هستند، هر دو ورودی می‌باشند. لذا در این مورد دو متغیر ورودی سود پرداختی و میزان سرمایه به‌عنوان مهم‌ترین متغیرهای ورودی و متغیر درآمد کل که بعد از این دو متغیر مهم‌ترین محسوب می‌گردد، به‌عنوان متغیر خروجی انتخاب شدند.

جدول ۹- متغیرهای انتخاب شده

متغیرهای ورودی	متغیرهای خروجی
سود پرداختی	درآمد کل
میزان سرمایه	

سپس نمرات کارایی واحدها با استفاده از فرم BCC ورودی‌گرا و با به‌کارگیری این شاخص‌ها محاسبه و در جدول ۱۰ آمده‌است.

جدول ۱۰- نمرات کارایی واحدها با استفاده از متغیرهای انتخاب شده

واحد تصمیم‌گیرنده	کارایی	واحد تصمیم‌گیرنده	کارایی
A1	۱,۰۰۰	A19	۰,۷۸۶
A2	۱,۰۰۰	A20	۱,۰۰۰
A3	۰,۹۱۵	A21	۰,۹۹۸
A4	۰,۹۲۷	A22	۰,۸۷۱
A5	۰,۹۴۴	A23	۱,۰۰۰
A6	۰,۹۷۶	A24	۰,۹۲۰
A7	۱,۰۰۰	A25	۰,۹۸۲
A8	۰,۹۵۵	A26	۱,۰۰۰
A9	۰,۹۴۵	A27	۱,۰۰۰
A10	۰,۸۹۳	A28	۰,۹۹۶
A11	۰,۸۴۸	A29	۰,۹۲۴

A12	۰,۸۵۶	A30	۰,۹۵۳
A13	۱,۰۰۰	A31	۱,۰۰۰
A14	۱,۰۰۰	A32	۰,۸۹۵
A15	۰,۹۷۹	A33	۰,۹۸۶
A16	۰,۸۷۵	A34	۱,۰۰۰
A17	۰,۹۶۸	A35	۰,۹۵۰
A18	۰,۹۸۲	A36	۰,۸۷۵

### نتیجه‌گیری

آنالیز تشخیصی روشی است که هدف آن پیدا کردن ترکیبی خطی از متغیرهایی مستقل است به گونه‌ای که این ترکیب از نمرات خام متغیرهای مستقل بهترین پیش‌بینی از تعلق واحدهای مورد بررسی به گروه‌ها را ممکن سازند.

در این پژوهش آنالیز تشخیصی با هدف ارزیابی اهمیت نسبی متغیرهای مستقل در طبقه‌بندی واحدها به دو گروه واحدهای کارا و ناکارا به کار گرفته شده است. هرچه این ضرایب بزرگ‌تر باشند، میزان تأثیر متغیر مستقل مربوطه در پیش‌بینی عضویت واحدها در دو گروه کارا و ناکارا بیش‌تر است. بنابراین از بین متغیرهای اولیه متغیرهایی که دارای بزرگ‌ترین ضرایب بودند، به عنوان مهم‌ترین متغیرها از بین متغیرهای اولیه انتخاب شدند. به این ترتیب از بین متغیرهای اولیه، متغیرهای ورودی سود پرداختی، میزان سرمایه و متغیر خروجی درآمد کل به عنوان مهم‌ترین متغیرهای تأثیرگذار در نمرات کارایی واحدها انتخاب شدند. انتخاب متغیرهای ورودی و خروجی که تأثیر فراوانی در تعیین کارایی و ناکارایی واحدهای تحت ارزیابی دارند، از اهمیت زیادی برخوردار است. با توجه به کاهش قدرت تشخیص تحلیل پوششی داده‌ها با افزایش تعداد ورودی‌ها و خروجی‌ها، انتخاب روشمند متغیرها راهی برای حفظ قدرت تشخیص و توان این روش در تحلیل کارایی واحدهایی است که با شاخص‌های متعدد ورودی و خروجی شناخته می‌شوند.

## منابع

جهانشاهلو غلامرضا، نیکومرام، هاشم و حسینزاده لطفی، فرهاد، تحلیل پوششی داده‌ها و کاربردهای آن، آثار نفیس، ۱۳۹۱.  
منصورفر، کریم، (۱۳۸۵)، روش‌های پیشرفته آماری همراه با برنامه‌های کامپیوتری، تهران: انتشارات دانشگاه تهران.

Banker, R.D., (1993), "Maximum likelihood, consistency and data envelopment analysis: A statistical foundation". *Management Science* 39 (10), pp. 1265–1273.

**Banker, R.D., (1996)**, "Hypothesis tests using data envelopment analysis". *Journal of Productivity Analysis* 7, pp. 139–159.

Berger, A., Humphrey, D., (1997), "Efficiency of financial institutions: International survey and directions for future research", *European Journal of Operational Research*, 98, pp. 175–212.

Boussofiene, A., Dyson, R.G., Thanassoulis, E., (1991), "Applied data envelopment analysis". *European Journal of Operational Research* 52, pp. 1–15.

Charnes, A., Cooper, W. W., Rhodes, E., (1978), "Measuring the efficiency of decision making units", *Eur. J. Oper. Res.* 2, pp. 429-444.

Charnes, A., Banker, R. D., Cooper, W.W., (1984), "Some Models for estimating technical and scale inefficiencies in Data Envelopment Analysis", *Management science* 30(9), pp. 1078-1092.

Farrell, M. J., (1957), "The measurement of productive efficiency", *Journal of the Royal Statistical Society Series a*, 120 (III), pp. 253-281.

Edirisinghe, N.C.P. & Zhang, X., (2010), Input/output selection in DEA under expert information, with application to financial markets, *European Journal of Operational Research*, Vol. 207 (3), pp. 1669–1678.

Golany, B., Roll, Y., (1989), "An application procedure for DEA". *OMEGA* 17 (3), pp. 237–250.

Jenkins, L., Anderson, M., (2003), "A multivariate statistical approach to reducing the number of variables in data envelopment analysis". *European Journal of Operational Research* 147, pp. 51–61.

Kittelson, S.A.C., (1993), *Stepwise DEA: Choosing variables for measuring technical efficiency in Norwegian electricity distribution*", Memorandum No. 06/93, Department of Economics, University of Oslo, Norway.



Klimberg, R., Puddicombe, M., (1995), "A multiple objective approach to data envelopment analysis", working paper 95-105, School of Management, Boston University, MA.

Nunamaker, T.R., (1985), Using data envelopment analysis to measure the efficiency of non-profit organizations: A critical evaluation. *Managerial and Decision Economics* 6 (1), pp. 50–58.

Pinto, U., Maheshwar, B., Shrestha, S. and Morris, C., (2012), Modeling eutrophication and microbial risks in peri-urban river systems using discriminant function analysis, *water research* 46, 6476- 6488.

Pardoe, Iain and et all, (2007), graphical tools for quadratic discriminant analysis, American statistical association and the American society for quality, *Technometrics*, may, vol (49), no. 2.

Salinas-Jimenez, J., Smith, P., (1996), "Data envelopment analysis applied to quality in primary health care". *Annals of Operations Research* 67, pp. 141–161.

Tabachnick, B. G. and Fidell, L. S., (2011), *Using Multivariate Statistics*, Fifth Edition, Omegatype Typogmph, Inc.

Tone, K., Cooper, W. W., Seiford, L. M., (2010), "Data Envelopment Analysis: A Comprehensive Text with Models", *Applications, References and DEA-Solver Software*. 2nd Edition, New York: Springer.

Warner, R. M., (2013), *Applied Statistics: From Bivariate Through Multivariate Techniques*, SAGE Publications.

Yan, L., Gongbing, B. and Liang, L., (2012), Input/output indicator selection for DEA efficiency evaluation: An empirical study of Chinese commercial banks, *Expert Systems with Applications* 39 (1), pp. 1118-1123.