Muthanna Journal of Administrative and Economics Sciences
مجلة المثنى للعلوم الادارية والاقتصادية
المجلات الاكاديمية العراقية

Using differentiation analysis method for classification and forecasting: An applied study in the iraq stock exchange

Sarah Jabr Hussein*

Dhi Qar University / Department of Control and Internal Audit.

Abstract

Discriminant analysis is one of the most common multivariate analysis methods. This analysis has excellent uses in different statistical studies, especially those concerned with classification and forecasting, because of its excellent ability to distinguish between data and classify new observations according to the groups to which they belong based on a particular function called the Discriminant Function. The main objective of this work is to investigate the differentiation analysis for classification and forecasting in financial analysis. In addition, this work aims to use the discriminatory function of classification, forecast the financial impact of a sample of Iraqi banks, and measure the classification error. Also, this work has set out the basic concepts of the classification processes and the conditions and assumptions to be provided for the application of the discriminatory method of analysis. Finally, this study has concluded that the proposed discriminatory function is highly effective in classifying and predicting with an appropriate classification error.

Information

Received: 2/5/2023
Revised: 24/5/2023
Accepted: 11/6/ 2023
Published: 30/9/2023

Keywords:
Discriminant analysis
financial ratios
Discriminant Functions
classification
prediction

أستخدام أسلوب تحليل التمايز في التصنيف والتنبؤ: دراسة تطبيقية في سوق العراق للأوراق المالية

سارة جبر حسين*

جامعة ذي قار / قسم الرقابة والتدقيق الداخلي

الملخص

يعتبر تحليل التمايز (Discriminant Analysis) واحد من أساليب التحليل الاحصائي متعدد المتغيرات (Multivariate Analysis) المشهورة وله استعمالات كبيرة في مختلف الدراسات الإحصائية خاصة تلك الدراسات التي تهتم بالتصنيف والتنبؤ لما له من قدرة كبيرة على التمييز بين البيانات وتصنيف المشاهدات الجديدة حسب المجموعات التي تنتمي اليها بالاعتماد على دالة خاصة تسمى الدالة التمييزية (Discriminant Function). كانت المشكلة الرئيسية التي واجهت الباحث هو كيفية استخدام تحليل التمايز للتصنيف والتنبؤ في مجال التحليل المالي. ان الهدف الأساسي من الدراسة هو استخدام الدالة التمييزية للتصنيف والتنبؤ بالتعثر المالي لعينة من المصارف العراقية وقياس خطأ التصنيف. كذلك عرض المفاهيم الأساسية لعمليات التصنيف والشروط والافتراضات الواجب توفرها لتطبيق أسلوب التحليل التمييزي. توصلت الدراسة الى ان الدالة التمييزية المقترحة ذات فاعلية عالية في التصنيف والتنبؤ وبخطأ تصنيف مناسب.

الكلمات المفتاحية: تحليل التمايز ،التعثر المالي ،النسب المالية ،التصنيف ،التنبؤ.

المقدمة

يعتبر التعثر المالي ظاهرة خطيرة تتعرض لها المؤسسات المالية والمصارف وقد اهتم المختصون منذ زمن بعيد بوضع الخطط والدراسات للسيطرة على هذه المشكلة المالية. لقد تم اقتراح نماذج مختلفة للتنبؤ بالتعثر المالي مثل نماذج (Kida,Altman,Sherrod,…) وغيرها. ان من اهم الأساليب الإحصائية التي استخدمت في مجال التحليل المالي هو أسلوب التحليل التمييزي وهو أحد أساليب التحليل الإحصائي متعدد المتغيرات وله استخدام واسع في التصنيف والتمييز والتنبؤ وغيرها، وهو أسلوب إحصائي يهتم بدراسة العلاقة بين المتغير التابع ومجموعة البيانات المستقلة. لقد بين (Maharaj, 2014) بأن الغرض من التحليل التمييزي هو تحديد المتغيرات التي تؤثر على عملية التمييز بين مجموعتين أو أكثر، وبناء قاعدة تمييزية للتنبؤ بانتماء عناصر جديدة إلى إحدى المجموعات بناءً على المعلومات الأولية المتاحة في البيانات

* Corresponding author : E-mail addresses : sarahjebur@utq.edu.iq.
2023 AL – Muthanna University . DOI:10.52113/6/2023-13-3/16-24

قيد الدراسة. يعتمد اسلوب التحليل التمييزي بشكل أساس على تكوين الدالة التمييزية التي هي عبارة عن تركيبة خطية بين المتغير التابع والمتغيرات المستقلة والتي تعظم الفروق بين متوسطات المجموعات قيد الدراسة كذلك تقليل التشابه في أخطاء التصنيف. يهدف التحليل التمييزي الى التحقق من وجود فروق ذات دلالة إحصائية بين المجموعات ولعدد من المتغيرات كذلك تحديد المتغيرات التي تسهم بشكل قوي في عمليات التصنيف وتقييم دقة التصنيف.

يشترط أسلوب التحليل التمييزي بان تتوزع البيانات وفق التوزيع الطبيعي كذلك تحقق فرضية تجانس التباين بين المجاميع قيد الدراسة واختيار عينة الدراسة بشكل عشوائي. وحسب (Feinberg, 2010)، فأن التحليل التمييزي يمكن أن يحل مشاكل التصنيف التي تنطوي على متغيرات تابعة فئوية وهو أداة جيدة للتنبؤ، لذلك فقد تم تطبيقه من قبل العديد من الباحثين في مختلف المجالات مثل الأعمال التجارية والطبية والتعليم والبيئة وعلم الاجتماع والتمويل وما إلى ذلك.

كانت المشكلة الأساسية التي قام على أساسها البحث هو عدم دقة نماذج التنبؤ العالمية الجاهزة عند تطبيقها على بيانات سوق العراق للأوراق المالية. لذلك كان الهدف الأساس من الدراسة هو بناء نموذج احصائي متعدد المتغيرات يناسب البيانات المالية في سوق العراق للأوراق المالية وهو استخدام أسلوب التحليل الاحصائي متعدد المتغيرات وبيان قدرة الدالة التمييزية على التمييز.

اما فرضية الدراسة فيمكن تمثيلها باختبار الفرضية الاتية:

H_0 : الدالة التمييزية ليس لها القدرة على التصنيف والتنبؤ وان المتوسطات متساوية أي ان:

$$\mu_1=\mu_2$$

H_1: الدالة التمييزية لها القدرة على التصنيف والتنبؤ وتوجد فروق معنوية بين المتوسطات أي ان:

$$\mu_1\neq\mu_2$$

اهتمت الدراسات السابقة بموضوع التحليل التمييزي منذ فترة بعيدة ترجع لبداية القرن العشرين حيث كانت البدايات البسيطة لفكرة التصنيف والتمييز بين البيانات بطريقة علمية، من الدراسات الأولى هي دراسة (Tildesley, 1921) التي استخدمت في مجال التصنيف والتمييز، كانت المشكلة الأساسية للدراسة هو كيفية تمييز الأشياء وتصنيفها على أساس قاعدة علمية ثابتة. قام (Tildesley, 1921) باستخدام طريقة معلمات المطابقة العصرية ل (K. Person) والتي تعتمد على تحديد الوزن والحجم للمشاهدات.

اقترح (P.C. Mahalanobis, 1936) فكرة جديدة للتمييز بطريقة احصائية حيث اقترح مقياس احصائي مهم يستخدم للتمييز والتصنيف بين مجتمعين وسمي (Mahalanobis Distance)، وهو عبارة عن إحصائة تقيس مربع المسافة او البعد بين متوسطي مجتمعين وكانت الصيغة الأساسية له كالاتي:

$$D^2 = (\bar{x}_1 - \bar{x}_2)' \Sigma^{-1} (\bar{x}_1 - \bar{x}_2) \quad (1)$$

حيث ان:

$\bar{x}_2, \bar{x}_1$: متجهي المتوسطات للمجموعتين الأولى والثانية على التوالي.

Σ: مصفوفة التباين والتباين المشترك (Var-Cov matrix)

تم تطوير هذا المقياس لاستخدامها في تصنيف العناصر الى احدى مجموعتين تتبع التوزيع الطبيعي ولها خصائص مشتركة.

بعد ذلك تم اقتراح الدالة التمييزية الخطية من قبل (R.Fisher, 1963) والتي تعتمد على مصفوفة التباين والتباين المشترك وهي تركيبة خطية بين المتغيرات المستقلة والمتغير المعتمد وصيغتها الرياضية هي:

$$Z = x' \Sigma^{-1} (\bar{x}_1 - \bar{x}_2) \quad (1)$$

حيث ان:

$\bar{x}_2, \bar{x}_1$: متجهي المتوسطات للمجموعتين.

Σ: مصفوفة ال (Var-Cov matrix)

بعد ذلك تم اقتراح دوال تمييزية جديدة مثل الدالة التمييزية التربيعية (Quadratic Discriminant Analysis) من قبل (Smith, 1946) والدالة اللوجستية (Logistic discriminant function) من قبل (Cox & Chambers, 1967) التي تستخدم عند انتهاك فرضية التوزيع الطبيعي للبيانات.

الإطار النّظري

ان تحللي التمايز (D.A) هو أحد أساليب التحليل الاحصائي متعدد المتغيرات والذي يستخدم للتصنيف والتنبؤ للمجموعات والتي تم توزيعها مسبقا على مجموعتين او أكثر من المجموعات المنفصلة ومعرفة خطوط التقاطع بين البيانات، ووضع قاعدة محددة لمعرفة انتماء أي عنصر جديد للمجموعة المناسبة. يعتمد التحليل التمييزي على بناء دالة تسمى الدالة التمييزية (Discriminant Function) وهي عبارة عن تركيبة خطية لعدد من المتغيرات المستقلة، هذه الدالة تقوم بتقليل التشابه في أخطاء التصنيف وزيادة التباينات بين المجتمعات. (بسيوني، 2021)

ان أساليب التحليل الاحصائي متعدد المتغيرات تأخذ بنظر الاعتبار العلاقات المتداخلة بين المتغيرات المستقلة المكونة وذلك من اجل الحصول على انموذج احصائي يبين العلاقات المتبادلة بين البيانات قيد الدراسة لذلك فان (D.A) ذو فائدة كبيرة لفاعليته في التمييز بين البيانات من خلال إيجاد علاقات خطية للمتغيرات التمييزية.

ان انموذج التحليل التمييزي يهدف للحصول على الدالة التمييزية (D.F.) التي تزيد من الاختلافات بين المتوسطات وفي نفس الوقت تقلل التشابه في الأخطاء وذلك باستخدام التوليفات الخطية بين المتغيرات. (Johnson & Wichern, 2002)

أولا. التحليل التمييزي الخطي Linear discriminant Analysis

اقترح (Ronald & Fisher, 1936) لأول مرة الدالة التمييزية الخطية (Linear Discriminant Function)، ان هذه الدالة هي الاسلوب الأفضل للتمايز في المسائل المعقدة وهي عبارة عن توليفة خطية وضعت كما في الصيغة الاتية:

$$Z = x' \Sigma^{-1} (\bar{x}_1 - \bar{x}_2) \quad (2)$$

حيث ان:

x': هو المبدل لمتجه المشاهدات الجديدة المراد تصنيفها.

$\bar{x}_2, \bar{x}_1$: متجهي المتوسطات للمجموعتين الأولى والثانية على التوالي.

Σ: مصفوفة التباين والتباين المشترك (Var-Cov matrix)

افترض (Fisher) بان يكون التباين متقارب للمجموعات وان البيانات تخضع للتوزيع الطبيعي، وان العلاقة يجب ان تكون خطية

بالنسبة للمعالم. تم استخدام هذه الدالة لدراسة عدد من النباتات اعتمادا على مجموعة من المتغيرات المرتبطة بحجم عينة (n=50) يراد تصنيف هذه المشاهدات الى مجموعتين، توصلت الدراسة الى ان هذه الدالة ذات قدرة عالية على التمييز بخطأ مناسب.

ثانياً. افتراضات التحليل التمييزي Assumptions of D.A.

هناك مجموعة من الافتراضات لاستخدام أسلوب التحليل التمييزي مثل:

1. المتغيرات المستقلة تتبع التوزيع الطبيعي المتعدد، اما إذا كان حجم العينة كبير فانه يقترب من التوزيع الطبيعي حسب نظرية الغاية المركزية. (Anderson, 1968)
2. المجموعات قيد الدراسة يجب ان تكون محددة بدون تقاطع، كذلك فان عددها (G>=2)، وذات حجم متجانس. (العلي، 2020)
3. مشاهدات العينة يتم اختيارها بطريقة مستقلة مثل طريقة المعاينة الطبقية بصورة عشوائية وان لا يقل عدد المشاهدات عن (30) حسب. (steven, 1996).
4. أن تكون المتغيرات ($X_1,X_2,\ldots,X_P$) مستقلة.
5. عدد المتغيرات يجب ان يكون اقل من عدد المشاهدات (P<n) وإذا كان عددها كبيرا فيجب تقليلها الى (5 %) من حجم العينة الكلي (n).
6. أن تكون مصفوفات (Var-Cov) متماثلة للمجموعات. (الرفاعي، 2006).
7. أن تكون(البواقي) خاضعة للتوزيع الطبيعي.
8. يجب ان تكون العينة ممثلة للمجتمع المدروس ويتم سحبها عشوائيا..
9. عدم وجود مشاهدات شواذ (Outliers) ضمن العينة.

ثالثاً. الغرض من استخدام تحليل التمايز The purpose of D.A.

يتم استخدام هذا الاسلوب لعدة أغراض اهمها:(Bentler, 1980)

1. تحديد الفروقات بين المجموعات للبيانات قيد الدراسة.
2. معرفة القاعدة المفضلة للتصنيف.
3. معرفة أي المتغيرات التنبؤية تساهم في التمييز والتصنيف.
4. اختزال المتغيرات ذات الارتباطات الضعيفة.
5. تصنيف المشاهدات ضمن المجموعات حسب قاعدة التصنيف.
6. التنبؤ بانتماء المشاهدات الجديدة باستخدام الدالة التمييزية.
7. معرفة دقة التصنيف بالاعتماد على خطأ التصنيف.

رابعاً. جانب من استخدامات تحليل التمايز Application of D.A.

ذكرت الدراسات السابقة استخدام هذا الأسلوب في عدة مجالات نذكر منها:

1. **مجال التعليم:** تم تطبيق التحليل التمييزي لتوزيع الطلبة على الاختصاصات العلمية حسب قدرتهم الذهنية والبدنية ومستوى الذكاء وكفاءة الطلبة.
2. **المجال الزراعي والبيولوجي:** تم تطبيق اسلوب التحليل التمييزي لتصنيف بعض أنواع النباتات الزراعية النادرة وبعض الكائنات الحية وتحديد السلالة او الصنف الذي تنتمي له حسب خصائص معينة.
3. **الجانب الطبي:** يتم استخدام أسلوب التحليل التمييزي في الجانب الطبي عند تصنيف حالة المريض حسب الخطورة.
4. **استخدامه للتنبؤ بالتعثر المالي:** استخدمت الدراسات السابقة مثل (Altman 1968, Kida , Sherrod) لأسلوب التحليل التمييزي لغرض التنبؤ بالتعثر المالي وتصنيف الشركات او المؤسسات المالية الى متعثرة وغير متعثرة بالاعتماد على المؤشرات المالية التي توفرها القوائم المالية لتلك المؤسسات والابتعاد عن مشاكل الإفلاس التي قد تحدث في المستقبل، وهو الجانب التطبيقي للبحث.

خامساً. خطوات تطبيق التحليل التمييزي The steps of D.A.

اهم الخطوات لتطبيق التحليل التمييزي هي:

1. معرفة نوع التحليل التمييزي وحسب طبيعة الدراسة.
2. معرفة المتغيرات التمييزية المناسبة الداخلة في النموذج.
3. عدد المجموعات (G) يجب ان يكون ($G\geq2$).
4. اختبار تجانس التباين للبيانات للتأكد من تحقق الافتراضات والشروط اللازمة لإجراء التحليل التمييزي.
5. تقدير المعاملات التمييزية.
6. تصنيف العناصر الجديدة حسب الدالة التمييزية ونقطة الفصل.

سادساً. تطبيق التحليل التمييزي للتنبؤ بالتعثر المالي

يعتبر التعثر او الفشل المالي ظاهرة خطيرة تتعرض لها المصارف والشركات وقد اهتمت البحوث لوضع حلول مناسبة للحد من هذه المشكلة وعدم الوقوع فيها. هنالك مفاهيم عديدة للتعثر المالي ولكن بالمفهوم العام هو عجز الشركة او المصرف على تسديد التزاماتها المالية حين استحقاقها حسب (Beaver, 1966). هناك تعريف اخر للتعثر المالي من قبل (Blum, 1974) هو سلسلة المراحل التي تؤدي بالنهاية إلى عدم قدرة الشركة على إيفاء الديون والالتزامات المالية التي عليها للأفراد او الشركات الأخرى. تبدأ الخطوات الأولى في انخفاض طلب المنتجات ثم ضعف الكفاءة وطرق الإنتاج ثم تتطور لتصل الى ضعف معدلات الدوران ونسب التمويل ثم ضعف الهيكل المالي للشركة ثم تنتهي أخيرا بالإفلاس إذا ما تم تدارك المشكلة ووضع الحلول المناسبة لها.

الجانب العملي

أولاً. البيانات الحقيقة Real data

تضمن الجانب العملي لهذه الدراسة على جمع بيانات لعينة عشوائية مكونة من (13) مصارف من المصارف المسجلة في سوق العراق للأوراق المالية الأكثر تداولا للفترة (2017-2010) حيث ان عدد المصارف الكلي بلغ (48) مصرف، اما المصارف التي شملتها عينة الدراسة فهي كل من (المصرف الأهلي، مصرف اشور، مصرف بابل، مصرف بغداد، مصرف ايلاف، مصرف الائتمان، مصرف الاستثمار، المصرف العراقي، مصرف الخليج، مصرف المنصور، مصرف سومر، المصرف التجاري والمصرف الوطني). تم اختيار (28) نسبة مالية في هذه دراسة وجمع البيانات عنها وتم اعتبارها متغيرات توضيحية فكان حجم العينة (104) مشاهدة.

ثانيا. تعريف المتغيرات Definition of Variables

اعتمد الجانب التطبيقي على جمع المالية المهمة في سوق العراق للأوراق المالية لعينة الدراسة. حيث بلغ عدد المؤشرات المالية (28) نسبة تم اعتبارها كمتغيرات توضيحية للدراسة وهي كل من:

1.نسبة العائد على السهم (X_1)
2.نسبة الملكية (X_2)
3.مكرر الأرباح (X_3)
4.نسبة التداول (X_4)
5.القيمة الدفترية (X_5)
6.سعر الاغلاق السنوي (X_6)
7.معدل السعر السنوي (X_7)
8.اعلى سعر نقد (X_8)
9.أدنى سعر نقد (X_9)
10. نسبة كفاية راس المال (X_{10})
11. نسبة السيولة (X_{11})
12. نسبة التشغيل (X_{12})
13. نسب الربحية وتشمل:
أ. نسبة صافي الأرباح الى راس المال (X_{13})
ب. نسبة صافي الدخل الى حقوق المساهمين (X_{14})
14. معدلات الدوران وتشمل:
أ. معدل دوران اجمالي الأصول (X_{15})
ب. معدل دوران الأصول المتداولة (X_{16})
15. نسب كلف النشاط وتشمل:
أ. نسبة الإيرادات الى اجمالي الودائع (X_{17})
ب. نسبة المصروفات الى الإيرادات (X_{18})
ج. نسبة المصروفات الى اجمالي الودائع (X_{19})
16. نسب الرافعة المالية وتشمل:
أ. نسبة الودائع الى (رأس المال + الاحتياطي)
ب. نسبة (رأس المال + الاحتياطي) الى الموجودات(X_{21})
17. القيمة السوقية (X_{22})
18. نسب اخرى وتشمل:
أ. نسبة السيولة (X_{23}) = (صافي راس المال العامل / الموجودات) *100
ب. نسبة السيولة (X_{24}) = (الأصول المتداولة / اجمالي الموجودات) *100
ج. مؤشر الرفع (X_{25}) = (حقوق المساهمين / الموجودات) *100
د. نسبة الربحية (X_{26}) = (صافي الربح قبل الضريبة/ اجمالي الموجودات) *100
ه. نسبة الرفع (X27) = (اجمالي الموجودات/ اجمالي المطلوبات) *100
و. نسبة الرفع (X_{28}) = (حقوق الملكية/ الموجودات الثابتة) *100

تم حساب جميع النسب أعلاه وكان المصدر الأساس للمؤشرات هو القوائم والنسب المالية لفترة الدراسة المنشورة في سوق العراق للأوراق المالية المرفوعة على الموقع الالكتروني (www.isx-iq.net).

لذلك فان المعادلة التمييزية للدراسة سوف تكون بالصيغة الاتية:

$$Z = \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_{28} X_{28} + \varepsilon \quad (3)$$

حيث ان $\beta_1, \beta_2, \dots, \beta_P$هي المعاملات التمييزية.
(p) هو عدد المتغيرات المستقلة الذي يساوي (28) في هذه الدراسة.
(ε) يمثل خطأ التصنيف للنموذج التمييزي.

ثالثاً. التحليل الاحصائي للبيانات Statistical analysis of data

تم الحصول على نتائج التحليل الاحصائي للمتغيرات المكونة لدالة التحليل التمييزي الخطية (LDA) من خلال استخدام لغة البرمجة (R. Program) من خلال الحزمة الاحصائية (klaR). حيث ان عينة المصارف المتعثرة وغير المتعثرة ينتميان الى مجتمعين معينين. ومن ثم تم تنظيم هذه النتائج كمدخلات لدالة التحليل التمييزي (LDA) في الحزمة الاحصائية (klaR) وحسب البيانات السنوية لكل مصرف. حيث تم ادخال ثمان عينات لكل مصرف من المصارف الـ وبحجم (104) مشاهدة لإجمالي المصارف قيد الدراسة. ان تصنيف حالات التعثر وعدم التعثر لكل مصرف حسب السنوات كان وفق معيار الأغلبية. حيث أظهرت النتائج وجود (48) حالة تعثر وباحتمال بلغ (0.46) و (56) حالة عدم تعثر وباحتمال مختلف (0.54)، كذلك فانه يوجد مصرف ما قد صنف اجمالا (خلال 8 سنوات) بانه متعثر لكن في سنة ما من تلك السنوات او بما لا يزيد عن (3) سنوات كان أداءه مميزاً فتعد تلك السنوات حالات عدم تعثر للمصرف والعكس صحيح. ويمكن ايجاز نتائج التحليل الاحصائي للنموذج التمييزي الخطي (LDA) كالاتي:

1.معرفة المتغيرات المهمة للدالة التمييزية

من اجل الحصول على أفضل نموذج تمييزي يجب اختيار المتغيرات المؤثرة والتي تؤثر على عمليات التصنيف ويتم هذا الاجراء باستخدام انحدار الخطوات الامامي (Stepwise forward selection) وكذلك اختبار (Wilks' Lambda) الذي تكون احصائته محصورة بين ($0 \leq \Lambda \leq 1$) فكلما اقتربت من (1) يكون المتغير أكثر تأثيرا في النموذج. ثم استخدام اختبار ((F.TEST Statistics ومعيار P.value)) عند مستوى دلالة (5%) لتحديد النسب المعنوية، فاذا كانت قيمة (P.value) لتلك النسبة اقل من (0.05) فان تلك النسبة من المتغيرات المهمة والتي تدخل في النموذج التمييزي كما في الجدول (1).

الجدول (1) تحديد المتغيرات المهمة باستخدام اختبار (Wilks' Lambda) واختبار (F) وقيمة (P.value)

المتغيرات المهمة	Wilks. Lambda Statistics	F. Statistics	P.value
X_{24}	0.8648150	15.944250	0.0001230
X_{19}	0.7935770	13.135910	8.5E-060
X_{28}	0.7583980	10.618950	4.02E-060
X_{17}	0.7264030	9.3219730	1.95E-060
X_{21}	0.7059570	8.1637380	1.79E-060

بينت النتائج في الجدول (1) بان هناك خمس متغيرات تعد من المتغيرات المهمة في المعادلة التمييزية حيث كانت اعلى قيمة لإحصائة (Wilks' Lambda) للمتغير (X_{24}) فقد بلغت (0.864815) وبذلك يعد من أكثر النسب المالية تأثيرا في النموذج الرياضي ثم تأتي بعده بقية النسب المالية من حيث الأهمية.

النموذج الرياضي يكون حسب المتغيرات الداخلة فيه من حيث الاهمية وفق الصيغة الاتية:

$$Y = \beta_{17}X_{17} + \beta_{19}X_{19} + \beta_{21}X_{21} + \beta_{24}X_{24} + \beta_{28}X_{28} \qquad (4)$$

حيث ان:

$(\beta_{28}, \beta_{24}, \beta_{21}, \beta_{19}, \beta_{17})$ هي المعاملات التمييزية للمتغيرات المختارة

اما النسب المهمة الداخلة في النموذج فهي:

(X_{17}): نسبة كلف النشاط

(X_{19}): نسبة المصروفات الى اجمالي الودائع

(X_{21}): نسبة الرافعة المالية

(X_{24}): مؤشر السيولة

(X_{28}): مؤشر الرفع

2.متوسطات المتغيرات المهمة والمعاملات التمييزية

بعد تحديد النسب المالية المهمة التي تؤثر في عمليات التصنيف لأسلوب التحليل التمييزي تم حساب متوسطات تلك النسب المالية للمصارف المتعثرة وغير المتعثرة كذلك المعاملات التمييزية للنموذج التنبؤي وحسب الجدول (2):

جدول (2) المتوسطات والمعاملات التمييزية للنسب ال المهمة للمصارف المتعثرة وغير المتعثرة والمعاملات التمييزية لنتائج التحليل التمييزي

النسب المالية المهمة	متوسط المؤشرات للمصارف المتعثرة	متوسط المؤشرات للمصارف غير المتعثرة	المعاملات التميزية
X_{24}	0.92	0.96	13.61
X_{19}	0.08	0.095	6.25
X_{28}	11.93	41.5	0.01
X_{17}	0.519	0.43	0.37
X_{21}	0.49	0.46	-2.04

حسب الجدول أعلاه تم تقدير المعاملات التمييزية والمتوسطات للنسب المالية المهمة التي هي المتغيرات التوضيحية. الان تم تكوين دالة تمييزية يمكن ان تستخدم للتنبؤ بالتصنيف للمشاهدات الجديدة وقد وضعت المعادلة التنبؤية حسب الصيغة ادناه:

$$y = 0.37x_{17} + 6.25x_{19} - 2.04x_{21} + 13.61x_{24} + 0.01x_{28}$$

من الواضح بان نسبة السيولة المالية وهي (X_{24}) تمتلك أكبر معامل حيث بلغت قيمته (13.61) لذلك فإنها أكثر المتغيرات تأثيرا في القيمة التنبؤية حيث يتناسب بشكل طردي كذلك فانه يمتلك أكبر قيمة لإحصائة (Wilks' Lambda) اما المتغير (X_{21}) الذي يمثل نسبة الرافعة المالية فهو يمتلك معامل تمييزي ذو قيمة سالبة وهي (-2.04) لذلك فان تأثيره على القيمة التنبؤية يتناسب بشكل عكسي.

3.والقيم التنبؤية Cutter point

ان نقطة الفصل هي المرحلة الأهم للتنبؤ في تصنيف العناصر الجديدة الى أحد المجموعتين (المتعثرة وغير المتعثرة) ويمكن حسابها وفق الصيغة الرياضية الاتية:

$$L = \frac{\bar{y}_1 + \bar{y}_2}{2} \qquad (5)$$

حيث ان ($\bar{y}_1$,$\bar{y}_2$) يمثل كل منهما المتوسط العام للمجموعتين الأولى والثانية ويمكن حسابهما وفق الصيغتين المعرفتان كالاتي:

$$\bar{y}_1 = \bar{x}_1' s^{-1}(\bar{x}_1 - \bar{x}_2) \qquad (6)$$

$$\bar{y}_2 = \bar{x}_2' s^{-1}(\bar{x}_1 - \bar{x}_2) \qquad (7)$$

وقد تم حساب القيم التنبؤية لجميع المصارف حسب السنوات كما في الجدول (3):

جدول (3) تصنيف المصارف الى مجموعتين متعثرة وغير متعثرة حسب السنوات بالاعتماد على القيم التمييزية

المصرف	السنة	القيم التمييزية	التصنيف	المصرف	السنة	القيم التمييزية	التصنيف	المصرف	السنة	القيم التمييزية	التصنيف
1. المصرف بابل	2010	-0.019	S	5. المصرف المنصور	2010	0.21	S	9. المصرف الائتمان	2010	1.689	S
	2011	-0.548	S		2011	0.109	S		2011	2.137	S
	2012	0.008	S		2012	-0.092	S		2012	0.978	S
	2013	0.181	S		2013	0.415	S		2013	1.559	S
	2014	-0.023	S		2014	0.457	S		2014	2.069	S
	2015	-0.051	S		2015	0.19	S		2015	1.947	S
	2016	-1.365	F		2016	-0.282	S		2016	0.79	S
	2017	-0.303	S		2017	0.161	S		2017	1.002	S

2. مصرف بغداد	2010	-0.682	F	6. مصرف سومر	2010	-0.244	S	10. مصرف الاستثمار	2010	0.051	S
	2011	-0.927	F		2011	-0.129	S		2011	-0.061	S
	2012	-0.916	F		2012	0.176	S		2012	0.087	S
	2013	-1.277	F		2013	1.219	S		2013	0.174	S
	2014	-2.746	F		2014	-0.476	S		2014	0.004	S
	2015	-2.299	F		2015	1.172	S		2015	-0.031	S
	2016	-1.697	F		2016	-1.121	F		2016	2.577	S
	2017	-0.386	S		2017	-0.644	F		2017	-0.549	S
3. مصرف ايلاف	2010	-0.662	F	7. المصرف الاهلي	2010	-0.323	S	11. مصرف الخليج	2010	-0.252	S
	2011	-0.116	S		2011	-0.778	F		2011	-0.506	S
	2012	-0.886	F		2012	-0.67	F		2012	-0.383	S
	2013	-1.518	F		2013	-1.218	F		2013	-0.227	S
	2014	-0.531	S		2014	3.446	S		2014	-0.398	S
	2015	-0.995	F		2015	-0.587	F		2015	-0.158	S
	2016	-1.632	F		2016	-0.945	F		2016	-0.457	S
	2017	-1.836	F		2017	-0.939	F		2017	-0.487	S
4. المصرف العراقي	2010	-1.081	F	8. مصرف أشور	2010	0.301	S	12. المصرف التجاري	2010	1.58	S
	2011	-1.607	F		2011	0.146	S		2011	1.219	S
	2012	-0.414	S		2012	0.163	S		2012	1.012	S
	2013	-0.161	S		2013	0.041	S		2013	1.368	S
	2014	-0.158	S		2014	0.041	S		2014	0.834	S
	2015	-1.399	F		2015	-0.344	S		2015	0.87	S
	2016	-0.687	F		2016	-1.259	F		2016	-0.139	S
	2017	-1.136	F		2017	-1.397	F		2017	0.039	S
13. المصرف الوطني	2010	-0.328	S								
	2011	0.807	S								
	2012	1.279	S								
	2013	1.325	S								
	2014	2.451	S								
	2015	2.259	S								
	2016	2.747	S								
	2017	2.191	S								

الجدول (3) يبين القيم التنبؤية وحالات النجاح والفشل لكل مصرف وحسب السنوات والتي تم حسابها بالاعتماد على الدالة التنبؤية حيث يوجد (8) حالات لكل مصرف وبعدد السنوات قيد الدراسة اذ تبين حالة المصرف في تلك السنة الدراسية لغرض إعطاء صورة واضحة عن حالة التعثر والنجاح (F,S) لكل سنة. حيث بلغت قيمة المتوسط للمصارف المتعثرة (-0.6680) و غير المتعثرة (0.4170) وعند التعويض في المعادلة (5) فان نقطة الفصل هي معدل هذين المتوسطين حيث بلغت:

$$L = \frac{\bar{y}_1+\bar{y}_2}{2} = \frac{0.4170+(-0.6680)}{2} = -0.1255$$

ان القيم التنبؤية التي اجتازت نقطة الفصل تعتبر غير متعثرة وبخلافه تكون متعثرة أي ان قرار التصنيف للعناصر الجديدة (x_0) يكون وفق الشرط الاتي:

تصنف القيمة الجديدة الى المجموعة الأولى عندما $L \leq Z_{x_0}$

تصنف القيمة الجديدة الى المجموعة الثانية عندما $L > Z_{x_0}$

حيث ان Z_{x_0} هي قيمة الدالة التمييزية عند تعويض القيمة التنبؤية الجديدة المراد تصنيفها.

4. نسبة خطأ التصنيف Misclassification

ان خطأ التصنيف (Misclassification) مؤشر ذو أهمية كبيرة لقياس كفاءة أداء الدالة التمييزية ويفضل الحصول على اقل خطأ تصنيف ممكن، اذ يمثل خطا التصنيف احتمالية تصنيف مشاهدة معينة الى مجموعة معينة مثلا (غير متعثرة) ولكنها في الحقيقة تنتمي الى مجموعة أخرى (متعثرة) والعكس صحيح.

عند ملاحظة الجدول السابق رقم (3) الذي يوضح القيم التنبؤية وحالات النجاح والفشل لكل مصرف بالنسبة للسنوات حيث يوجد ثمان حالات لكل مصرف بعدد سنوات الدراسة تبين حالة المصرف في تلك السنة لغرض إعطاء صورة تفصيلية عن حالة الفشل والنجاح (F,S).على سبيل المثال نلاحظ بان مصرف بابل قد استمرت سنواته بالنجاح (S)حتى سنة (2016) التي انتجت حالة فشل (F). مثال اخر

لنأخذ مصرف بغداد فان نتيجة التصنيف للسنوات من (2010) الى (2016) كانت متعثرة ولكن سنة (2017) أعطت حالة نجاح، فبالرغم من تصنيفه بانه متعثر الا انه يحوي سنة نجاح كذلك الحال في بقية المصارف توجد حالات تطابق بالتصنيف وحالات عدم تطابق لذلك يتولد لدينا ما يسمى بخطأ التصنيف والجدول رقم (4) يوضح حالات التعثر وعدمه التي ظهرت في نتائج التحليل الاحصائي للدالة التمييزية في جدول (3).

جدول (4) خطأ التصنيف للتنبؤ بالتعثر وعدم التعثر لعينة الدراسة

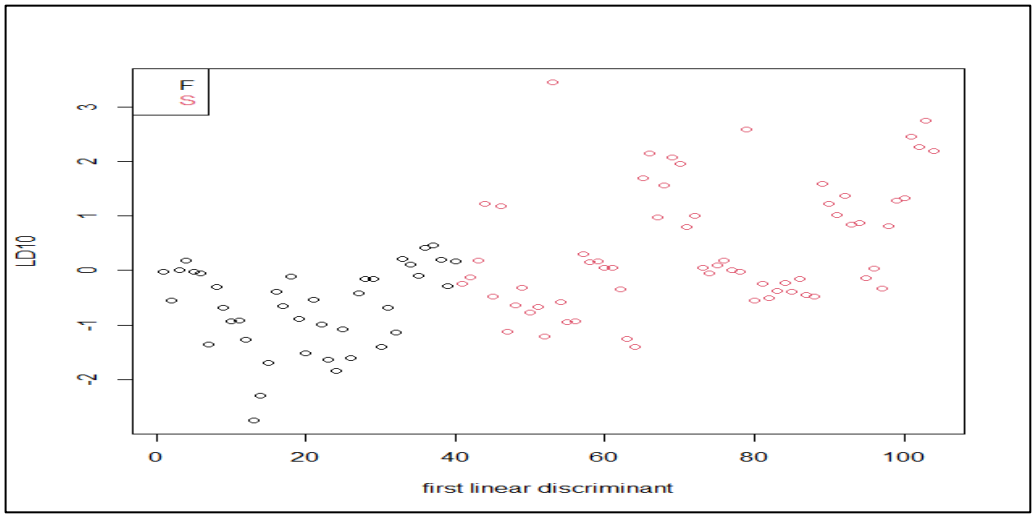
التنبؤ	الفرضية	
	تعثر	غير متعثر
تعثر	19	21
غير متعثر	10	54

ان قيمة خطأ التصنيف قد بلغت (0.30) وهي نسبة الحالات التي لم تتطابق مع نتائج التصنيف وهي نسبة جيدة ويمكن ان تعطي نتائج مقبولة في التصنيف والتنبؤ حيث بالمقابل تكون نسبة صحة التصنيف (0.70)، ومن اجل التأكد من التطابق بالتصنيف يتم جمع حالات التعثر التي تم تصنيفها متعثرة وهي متعثرة فعلا والتي تبلغ (19) حالة مع حالات عدم التعثر والتي تم تصنيفها بانها غير متعثرة أيضا وهي (54) ثم تقسيم الناتج على العدد الكلي لحجم العينة فتكون النتيجة:

(104/73=54+19 = 0.70)

نحصل على صحة التصنيف وهي (0.70) ثم نطرح من الواحد هذه النسبة نحصل على خطأ التصنيف الذي يبلغ (0.30).

الشكل (1) يوضح تقاطع القيم بعد التصنيف



الشكل أعلاه يوضح تقاطع المشاهدات الجديدة بعد التصنيف الى مجموعتين حيث ان المشاهدات التي باللون الاحمر على اليمين تمثل البيانات التي صنفت ضمن المصارف الغير متعثرة اما المشاهدات التي باللون الاسود فتشير الى البيانات التي صنفت الى المصارف المتعثرة ومن الواضح بان هناك تداخل بسيط بين البيانات عند حد معين هذه البيانات التي تشترك بين المجموعتين تعبر عن وجود خطأ تصنيفي قليل نسبيا مقارنة مع التصنيف الصحيح الذي يشكل الكمية الأكبر من البيانات.

اختبار (Wilks' Lambda) لقياس قدرة الدالة على التمييز

لبيان قدرة الدالة التميزية على التمييز بين المصارف وتحديد حالة التعثر تم استخدام اختبار (Wilks' Lambda). تم وضع فرضية لاختبار تساوي متوسطات المجموعتين حيث تشير فرضية العدم الى عدم وجود فروقات معنوية بين المتوسطات أي عدم قدرة الدالة على التمييز اما الفرضية البديلة فتشير الى وجود فروق معنوية بين المتوسطات أي قدرة الدالة على التمييز وكالاتي:

$$H_0: \mu_1 = \mu_2$$
$$H_1: \mu_1 \neq \mu_2$$

توصلت نتائج التحليل الاحصائي الى ان قيمة (Wilks' Lambda) تساوي (0.780). اما قيمة (P.value) فإنها تساوي (0.000150) وهذا مما يرجح رفض فرضية العدم وان الدالة التمييزية لها قدرة عالية على التمييز بين المجموعات قيد الدراسة.

الاستنتاجات والتوصيات

الاستنتاجات

توصلت الدراسة الى ان:

1. النموذج التمييزي المتمثل بالدالة التمييزية الخطية قد اعطى نتائج جيدة في التصنيف والتنبؤ حيث كان خطا التصنيف ما يقارب (3%) وهي نسبة خطا مسموح بها في الدراسات المتعلقة بالتحليل المالي.

2.ومن خلال ملاحظة اختبار قدرة الدالة اعلى التمييز فان احصاءة (Wilks' Lambda) كانت جدا عالية مما يعطي كفاءة عالية في التصنيف والتنبؤ.

3.كذلك فان قيمة (P.value) كانت منخفضة جدا مما يشير الى معنوية الفروق بين المتوسطات وان الدالة التمييزية ذات فاعلية قوية في التمييز.

التوصيات

لقد اوصت الدراسة بما يلي:

1. استخدام الأساليب الإحصائية المعتمدة مثل التحليل التمييزي في مجال التحليل المالي.
2. كذلك نوصي المؤسسات المالية بالتركيز على المتغيرات القوية التي اجتازت اختبارات الكفاءة في هذه الدراسة.

المصادر

بدر، دريد حسين. (2013). "استخدام بعض الطرائق التمييزية الحصينة لتشخيص امراض القلب في البصرة" مجلة دراسات البصرة، العدد (15) ,2013.

بسيوني، عبد الرحيم عوض عبد الخالق. (2021). " استخدام التحليل التمييزي في التصنيف والتنبؤ" دراسة تطبيقية، جامعة طنطا، مصر,2021.

الجهماني، عيسى. (2001). "استخدام النسب المالية للتنبؤ بتعثر المصارف", دراسة تطبيقية على القطاع المصرفي، مجلة الإدارة العامة، العدد (4)، الأردن, 2001.

الحسناوي، سالم صلال راهي. (2014). "التنبؤ بالتعثر المالي لشركات الاستثمار باستعمال النسب المالية لنموذج (Kida) وأثره على دراسة (EPS) ",دراسة في سوق العراق للأوراق المالية، مجلة المثنى للعلوم الإدارية والاقتصادية، المجلد (4)، العدد (9) ,2014.

حسين، محمد رضى عتاب. (1999). "استخدام الدالة المميزة التربيعية في تمييز أنماط الأرقام العربية " أطروحة دكتورا، كلية الإدارة والاقتصاد، جامعة بغداد,1999.

خوالدي، سليمة. (2017). "دور التحليل التمييزي في التنبؤ بالفشل المالي لعينة من المؤسسات الصغيرة والمتوسطة", رسالة ماجستير، كلية العلوم الاقتصادية والتجارية، ام البواقي، الجزائر,2017.

السليماني، مؤيد سلمان عباس. (1998). "استخدام الدالة المميزة لتشخيص حالات التهاب الأمعاء عند الأطفال", رسالة ماجستير في الإحصاء، كلية الإدارة والاقتصاد، جامعة بغداد,1998.

الكاتب، محمد أسامة. (2010). "الدالة المميزة الخطية في حالة أكثر من مجموعتين", مجلة التربية والعلوم، المجلد (23)، العدد (4)، المعهد التقني، نينوى,2010.

Anderson J.A. (1968). Constrained Discrimination between k Populations, Oxford university, journal of the royal statistical society, series B (Methodological) 31 (1) 123-129.

Chambers E. A. & Cox D.R. (1967). Discrimination between alternative binary response models, Imperial college, British.

Croux, C., & Dehon, C. (2001). Robust linear discriminant analysis using S-estimators. Canada. J. Statist., 29(3), 473–493. doi:10.2307/3316042.

D. J. Hand. (1983). A Comparison of Two Methods of Discriminant Analysis Applied to Binary Data, Biometrics, Vol. 39, No. 3 (Sep., 1983), pp. 683-694.

D. M. Titterington, G. D. Murray, L. S. Murray, D. J. Spiegelhalter, A. M. Skene, J. D. F. Habbema, G. J. Gelpke. (1981). Comparison of Discrimination Techniques Applied to a Complex Data Set of Head Injured Patients, Journal of the Royal Statistical Society. Series A (General), Vol. 144, No. 2 (1981), pp. 145-175.

Edward I. Altman. (1968). Financial ratios, discriminant analysis and the prediction of carport bankruptcy, The Journal of Finance, Vol. 23, No. 4 (Sep., 1968), pp. 589-609, America.

Feinberg, F. M. (2010). Discriminant analysis for marketing research applications. In N. S. Jagdish, & K. M. Naresh (Eds.) Wiley International Encyclopedia of Marketing (Vol. 2). New York, NY: John Wiley & Sons, Ltd doi:10.1002/9781444316568.

I. Akeyede, F.G. Ibi, D. T. Ailobhio. (2021). A Discriminant Analysis Procedure for Loan Application using Ranked Data, Asian journal for mathematic sciences, Vol 5 n 1 ISSN 2581-3463.

M.L. Tildesley. (1921). A first study of the Burmese skull, University College, London.

Maharaj, E. A., & Alonso, A. M. (2014). Discriminant analysis of multivariate time series: application to diagnosis based on ECG signals. Computational Statistics & Data Analysis, 70, 67-87. doi: 10.1016/j.csda.2013.09.006.

Michael Q. Zhang. (2000). Discriminant analysis and its application in DNA sequence motif recognition, Briefings in Bioinformatics, Volume 1, Issue 4, 1 November 2000, Pages 331–342.

Mufda J. Alrawashdeh, Taha Radwan and Khalid Abunawas. (2018). Performance of linear discriminant analysis using different robust methods, European Journal of pure and applied mathematics, Vol. 11, No. 1, 2018, 284-298.

P.C. Mahalanobis. (1936). On the generalized distance in statistics, India.

R.A. Fisher. (1936). The use of multiple measurements in taxonomic problem, the Annals of Eugenics, v. 7, p. 179-188 (1936) with permission of Cambridge University, London.

Randles, R. H., Broffitt, J. D., Ramberg, J. S., & Hogg, R. V. (1978). Generalized linear and quadratic discriminant functions using robust estimates. Journal of the American Statistical Association, 73(363), 564-568. doi: 10.2307/2286601.

Richard J. Beckman and Mark E. Johnson. (1981). A Ranking Procedure for Partial Discriminant Analysis, Journal of the American Statistical Association, Vol. 76, No. 375 (Sep., 1981), pp. 671- 675.

Ronald H. Randles, James D. Broffitt, John S. Ramberg & Robert V. Hogg. (1978). Generalized Linear and Quadratic Discriminant Functions Using Robust Estimates, Journal of the American Statistical Association, 73:363, 564-568.

Shelley B. Bull and Allan Donner. (1987). The Efficiency of Multinomial Logistic Regression Compared with Multiple Group, Journal of the American Statistical Association, Vol. 82, No. 400 (Dec., 1987), pp. 1118- 1122.

Smith, C.A. (1946). Some examples of discrimination, Annals of Eugene 13(1)272-382.