African Journal of Advanced Pure and Applied Sciences (AJAPAS)
Online ISSN: 2957-644X
Volume 4, Issue 1, January - March 2025
Page No: 520-533
Website: https://aaasjournals.com/index.php/ajapas/index
ISI 2024: 0.877 **SJIFactor 2024: 6.752** **معامل التأثير العربي: 1.62**

استخدام التحليل التمييزي في التصنيف
دراسة تطبيقية في تصنيف بعض الدول حسب النمو الاقتصادي

أ. الشتيوي إمحمد إمسيلخ [1]، أ. أحمد محمد سوادي [*2]
[1] قسم الاحصاء، كلية العلوم الزنتان، جامعة الزنتان، الزنتان، ليبيا
[2] قسم الرياضيات، كلية العلوم الأصابعة، جامعة غريان، الأصابعة، ليبيا

Using Discriminant Analysis in Classification: An Applied Study on Classifying Selected Countries by Economic Growth

Alshiteewi Amhimmid Imseelikh [1], Ahmed Mohamed Sawadi [2*]
[1] Department of Statistics, Faculty of Science, Zintan, University of Zintan, Zintan, Libya
[2] Department of Mathematics, Faculty of Science, Asabaa, University of Gharyan, Asabaa, Libya

[*]Corresponding author: ahmedswadi6@gmail.com

تاريخ الاستلام: 2025-02-13	تاريخ القبول: 2025-03-23	تاريخ النشر: 2025-03-27

الملخص

يهدف هذا البحث إلى دراسة قدرة التحليل التمييزي في التمييز بين الدول ذات النمو الاقتصادي المرتفع والدول ذات النمو الاقتصادي المنخفض استنادًا إلى ثلاث متغيرات مستقلة: معدل النمو الاقتصادي، نصيب الفرد من إجمالي الناتج المحلي، وإجمالي تكوين رأس المال.

تم تطبيق التحليل التمييزي الخطي باستخدام برنامج SPSS، حيث تم اختبار معنوية النموذج باستخدام اختبار Wilks' Lambda لتحديد مدى فعالية التحليل في تصنيف الدول بدقة. تم اختبار فرضية العدم التي تنص على عدم قدرة التحليل التمييزي على التمييز بين المجموعتين، وأظهرت النتائج أن قيمة الدلالة كانت أقل من 0.05 (Sig = 0.002)، مما يشير إلى وجود تمييز معنوي بين المجموعتين. كما تبين أن متغيرات "النمو الاقتصادي" و"نصيب الفرد من إجمالي الناتج المحلي" كان لها تأثير معنوي في عملية التصنيف، بينما لم يكن لمتغير "إجمالي تكوين رأس المال" تأثير معنوي. تمكن النموذج من تصنيف 82.4% من الحالات الأصلية بشكل صحيح، مما يعكس فعالية عالية للتحليل التمييزي. استند البحث إلى بيانات اقتصادية من قاعدة بيانات البنك الدولي لعام 2023، حيث تم تصنيف الدول إلى مجموعتين بناءً على متوسط معدل النمو الاقتصادي العالمي.

الكلمات المفتاحية: التحليل التمييزي، النمو الاقتصادي، المتغيرات.

Abstract

This research aims to study the ability of discriminant analysis to distinguish between countries with high economic growth and countries with low economic growth, based on three independent variables: economic growth rate, GDP per capita, and gross capital formation.

The linear discriminant analysis was applied using SPSS, and the significance of the model was tested using Wilks' Lambda to assess the effectiveness of the analysis in accurately classifying the countries. The null hypothesis, which states that discriminant analysis cannot differentiate between the two groups, was tested. The results showed that the significance value was less than 0.05 (Sig = 0.002), indicating a significant distinction between the two groups. It

was also found that the variables "economic growth" and "GDP per capita" had a significant effect on the classification process, while the variable "gross capital formation" did not have a significant effect. The model was able to correctly classify 82.4% of the original cases, reflecting the high effectiveness of discriminant analysis. The research relied on economic data from the World Bank database for 2023, with countries classified into two groups based on the global average economic growth rate.

Keywords: discriminant analysis, economic growth, variables.

مقدمة

يعد التحليل التمييزي أداةً إحصائيةً فعّالة تستخدم في تصنيف المجموعات بناءً على خصائصها المشتركة، مما يسهم في فهم الفروق بين المجموعات وتحليل العوامل المؤثرة فيها. في هذا البحث، تم تطبيق التحليل التمييزي لتصنيف الدول إلى مجموعتين، بناءً على معدلات النمو الاقتصادي، بهدف تقييم مدى دقة التحليل التمييزي في التصنيف وقدرته على تحديد الفروق بين المجموعات. وتبرز أهمية هذه الدراسة من خلال تسليط الضوء على استخدام التحليل التمييزي في المجالات الاقتصادية، واستكشاف إمكانياته في تحسين اتخاذ القرارات وتصنيف الظواهر المعقدة.

مشكلة البحث:

في ظل التطور الكبير في الأساليب الإحصائية المتقدمة لتحليل البيانات متعددة المتغيرات، يظهر تحليل التمييز الخطي كأحد الأدوات الفعّالة لتصنيف المجموعات بناءً على متغيرات متعددة. إلا أن هناك تحديات عديدة تتعلق باستخدام هذا التحليل، منها تحديد مدى دقة النموذج في تصنيف العينات الجديدة، وتحليل افتراضاته الأساسية، بناءً على ذلك تركز هذه الدراسة على اختبار قدرة التحليل التمييزي في التمييز بين الدول ذات النمو الاقتصادي المرتفع والمنخفض باستخدام المتغيرات الاقتصادية. لذلك، يطرح السؤال:

إلى أي مدى يمكن للتحليل التمييزي أن يساهم في تصنيف الدول بناءً على معدلات النمو الاقتصادي؟

أهمية البحث:

تنبع أهمية هذا البحث من كونه يُسهم في توظيف التحليل التمييزي كأداة إحصائية فعّالة في تصنيف الدول بناءً على أدائها الاقتصادي، مما يساعد في فهم الفروقات الجوهرية بين الدول ذات النمو الاقتصادي المرتفع وتلك ذات النمو المنخفض. كما يُظهر البحث كيف يمكن استخدام البيانات الاقتصادية المتوفرة لدعم التنبؤات والتصنيفات التي تساهم في رسم السياسات الاقتصادية والاستثمارية المستقبلية على المستويين الإقليمي والعالمي.

أهداف البحث:

يهدف البحث إلى:

1. دراسة مدى قدرة التحليل التمييزي على التمييز بين الدول ذات النمو الاقتصادي المرتفع والدول ذات النمو الاقتصادي المنخفض.
2. تحليل أثر ثلاث متغيرات اقتصادية مستقلة وهي: معدل النمو الاقتصادي، نصيب الفرد من الناتج المحلي الإجمالي، وإجمالي تكوين رأس المال، على عملية التصنيف.
3. قياس دقة النموذج التمييزي في تصنيف الدول وفقًا للفئة الاقتصادية الصحيحة، والتحقق من فعاليته.

فرضيات البحث:

الفرضية الصفرية (H_0): لا يمتلك التحليل التمييزي قدرة على التمييز بين الدول ذات النمو الاقتصادي المرتفع والدول ذات النمو الاقتصادي المنخفض بناءً على معدل النمو الاقتصادي.

متغيرات البحث:

المتغير التابع: (متغير نوعي) الذي يمثل المجموعة (نمو اقتصادي منخفض (0)، نمو اقتصادي مرتفع ((1)

المتغيرات المستقلة: إجمالي تكوين رأس المال (% من إجمالي الناتج المحلي)، نصيب الفرد من إجمالي الناتج المحلي (بالأسعار الجارية للدولار الأمريكي)، النمو الاقتصادي %.

مصادر البيانات:

تم الاعتماد على بيانات اقتصادية مستخلصة من قاعدة بيانات البنك الدولي لعام 2023، تم تصنيف الدول إلى مجموعتين: دول ذات نمو اقتصادي مرتفع ودول ذات نمو اقتصادي منخفض، وذلك بناءً على متوسط معدل النمو الاقتصادي العالمي، حيث تم تطبيق التحليل التمييزي الخطي باستخدام برنامج SPSS20 .

الدراسات السابقة:

1- دراسة (Hardigan, et al 2001):

هدفت هذه الدراسة إلى استخدام التحليل التمييزي للتنبؤ بقبول الطلبة في كلية الكيمياء، استنادًا إلى سبعة منبئات شملت: درجات الطلبة في البيولوجي، الكيمياء، الرياضيات، اللغة الإنجليزية، المعدل التراكمي، درجات اختبار PCAT، ورسائل التوصية. تم تطبيق التحليل على عينة من (486) طالبًا.وقد أظهرت نتائج الدراسة أن كلًا من: درجات البيولوجي، المعدل التراكمي، الدرجة الكلية لاختبار PCAT، درجات الرياضيات، الجزء الكمي، جزء استيعاب المقروء من اختبار PCAT، ورسائل التوصية، لها إسهام دال إحصائيًا في التنبؤ بقبول الطلبة. كما أشارت النتائج إلى أن الدالة التمييزية تمكنت من تصنيف حالات القبول بدقة بلغت 83.33%، مما يعكس فعالية التحليل التمييزي في التنبؤ بقرارات القبول الأكاديمي بناءً على مؤشرات متعددة.

2- دراسة (صالح، 2008):

هدفت هذه الدراسة إلى تصنيف بعض أمراض العيون إلى ثلاث مجموعات رئيسية بناءً على مجموعة من المتغيرات المصاحبة لكل مرض، وذلك باستخدام أسلوب التحليل التمييزي الخطّي (LDA). وقد تم سحب عينة عشوائية ممثلة من كل مجموعة على حدة، ومن ثم تطبيق الدالة التمييزية لاستخلاص الخصائص المميزة لكل مجموعة مرضية. أظهرت نتائج التحليل كفاءة الأسلوب المستخدم في التمييز بين المجموعات، مما يعكس قدرة المتغيرات المدخلة على تفسير الفروق بين أنواع أمراض العيون المدروسة.

3- دراسة (الشمراني، 2009):

استهدفت الدراسة توضيح كيفية استخدام التحليل التمييزي وتحليل التباين المتعدد في تحليل البيانات متعددة المتغيرات، سواء في حالة عامل واحد أو عاملين، مع تقييم كفاءة النموذج التمييزي واختبار قدرته التمييزية، إضافة إلى مقارنة نتائج التحليلين. استخدمت الدراسة المنهج الوصفي المقارن، وتم توليد البيانات باستخدام برنامج Minitab بناءً على معالم بيانات حقيقية. وأظهرت النتائج تقارب افتراضات التحليل التمييزي وتحليل التباين المتعدد، مع تفوق التحليل التمييزي في شمولية اختبار الافتراضات. كما أظهرت الدراسة أن قيمة ولكس لمدا وقيم مصفوفتَي معاملات الارتباط والتغاير متطابقة في كلا التحليلين، وأن تحليل التباين المتعدد يبرز الفروق بين المجموعات في المتغيرات التابعة قبل تطبيق التحليل التمييزي.

4- دراسة (Erimafa, 2009):

هدفت هذه الدراسة إلى استخدام التحليل التمييزي للتنبؤ بفئة الشهادة الجامعية المتاحة للطلاب الخريجين، وذلك من خلال تطوير دالة تمييز خطية تعتمد على متغيرات تنبؤية. أظهرت نتائج الدراسة نجاح الدالة في تصنيف فئات الشهادات بدقة بلغت 87.5% على العينة الأصلية، في حين ارتفعت النسبة إلى 88.2% عند تعميم النموذج على حالات جديدة غير معروفة العضوية، مما يدل على كفاءة الدالة في التصنيف والتنبؤ. كما أشارت الدراسة إلى أن أداء الدالة التمييزية في تصنيف الطلبة الجدد كان مماثلًا تقريبًا لأدائها مع العينة التاريخية، مما يعزز من صلاحية استخدام هذا النموذج كأداة إحصائية فعالة في تصنيف الطلاب وفق نوع الشهادة الجامعية المحتملة.

5- دراسة (النعيمي، 2019):

هدفت دراسة إلى الكشف عن عوامل التعثر الدراسي لدى طلبة الجامعات السعودية باستخدام التحليل التمييزي، حيث تم تحليل دور اختبار القدرات العامة والاختبار التحصيلي في التنبؤ بالتعثر. شملت العينة 31,082 طالباً وطالبة من ست جامعات سعودية، وأظهرت النتائج أن أداء المتعثرين كان أقل في كلا

الاختبارين، مع تفوق الأهمية النسبية للاختبار التحصيلي. كما تبين أن الدالة التمييزية كانت فعالة في التمييز بين المتعثرين وغير المتعثرين، مما يبرز أهمية تحسين معايير القبول الجامعي.

6- دراسة (نده، 2019):

هدفت هذه الدراسة إلى تصنيف المستوى المعيشي للأسر في محافظة اللاذقية إلى ثلاث فئات: منخفض، متوسط، وعالٍ، وذلك باستخدام التحليل التمييزي كأداة إحصائية لتحديد العوامل الأكثر تأثيرًا في هذا التصنيف. شملت المتغيرات المدروسة: الحالة العملية لرب الأسرة، المستوى التعليمي لرب الأسرة، الإنفاق الشهري للأسرة على الطعام واللباس، حجم الأسرة، مساحة المسكن، ونوع المسكن.

توصلت الدراسة إلى أن التحليل التمييزي كان فعالًا في تصنيف الأسر بدقة وفقًا لمستواها المعيشي، حيث أظهرت النتائج وجود أثر تصنيفي معنوي لكل من: الإنفاق الشهري على الطعام واللباس، حجم الأسرة، والحالة العملية لرب الأسرة. بينما لم يكن هناك أثر تصنيفي معنوي لكل من: المستوى التعليمي لرب الأسرة، مساحة المسكن، ونوع المسكن.

7- دراسة (Stella, 2019):

هدفت هذه الدراسة إلى استعراض التحليل التمييزي كأحد الأساليب الإحصائية متعددة المتغيرات المستخدمة في التنبؤ وتصنيف الأفراد إلى مجموعتين أو أكثر بناءً على مجموعة من المتغيرات المستقلة. تمثّل الهدف الأساسي للدراسة في بناء نموذج تنبؤي يُستخدم لتحديد عضوية الأفراد ضمن مجموعات محددة، وذلك من خلال دالة تمييزية أو مجموعة دوال في حال تعدد المجموعات.

وقد بيّنت نتائج الدراسة أن التحليل التمييزي يوفر أداة فعالة لتصنيف الأفراد بناءً على المتغيرات المدخلة، كما أظهرت أهمية معاملات الدالة التمييزية في تحديد مدى مساهمة كل متغير في التصنيف. كما أكدت الدراسة على إمكانية استخدام مصفوفة بنية العوامل لشرح العلاقات بين المتغيرات والدوال التمييزية بشكل أعمق، مما يُسهم في تقديم تفسيرات أكثر دقة للدوال الناتجة وتحسين جودة التصنيف.

8- دراسة (الشمراني، 2020):

هدفت هذه الدراسة إلى تصنيف بيانات الطلبة حسب الأداء الأكاديمي (مرتفع/منخفض) باستخدام التحليل العنقودي والتحليل التمييزي، وذلك على عينة مكونة من 62 طالبًا وطالبة. أظهرت نتائج تحليل العناقيد (K-Means) وجود مجموعتين مميزتين، كما أثبت التحليل التمييزي أهمية المتغيرين (X2) و(X3) في التصنيف بدلالة إحصائية قوية. بلغت نسبة الارتباط بين المتغيرات (0.770) وقيمة ولكس لمدا (0.408)، مما يدل على فعالية النموذج التمييزي، حيث بلغت دقة التصنيف 98.4%، وهو ما يؤكد كفاءة الأسلوبين في التمييز بين مستويات الأداء الأكاديمي.

المبحث الأول: التحليل التمييزي The Discriminant analysis

1- تمهيد:

يُعد التحليل الإحصائي من الأدوات الأساسية في البحث العلمي، إذ يُستخدم لتحليل البيانات الكمية واستخلاص النتائج بطريقة علمية. ومن بين أنواعه، يبرز التحليل متعدد المتغيرات الذي يهدف إلى دراسة العلاقة بين عدة متغيرات في الوقت ذاته، والذي يجعله مناسبًا لتحليل الظواهر المعقدة.

ويُعد التحليل التمييزي أحد أساليب التحليل متعدد المتغيرات الذي يُعد أداة فعّالة في تصنيف الوحدات الإحصائية إلى مجموعات، استنادًا إلى مجموعة من المتغيرات المستقلة، وتحديد العوامل الأكثر تأثيرًا في التمييز بين تلك المجموعات. ويُستخدم التحليل التمييزي بشكل واسع في مجالات متنوعة كالتعليم، الاقتصاد، والطب، نظرًا لدقته في التصنيف والتنبؤ.

2- التحليل التمييزي:

يهدف أسلوب التحليل التمييزي إلى التمييز بين مجموعتين أو أكثر من المشاهدات أو المفردات التي تم تحديدها مسبقاً، مع تخصيص المشاهدات أو المفردات الجديدة للمجموعة المناسبة التي تنتمي إليها.

ويعتمد التحليل التمييزي على وجود مجموعة من المشاهدات أو المفردات التي تشترك في مجموعة من الخصائص والصفات بدرجات مختلفة، مما يتيح إمكانية تحديد المجموعة التي تنتمي إليها أي من هذه المشاهدات أو المفردات، سواء كانت ضمن نطاق المجموعات أو خارجه. (عاشور وآخرون، 2024)

يهدف نموذج التحليل التمييزي إلى بناء دالة تمييزية (Discriminant Function - D.F) تعمل على زيادة الفروقات بين المتوسطات وتقليل التشابه في الأخطاء، من خلال استخدام تركيبات خطية للمتغيرات. (Johnson & Wichern, 2002)

3- أهداف التحليل التمييزي: (الزاوي، 2023)

1- أتصميم التوليفات الخطية للمتغيرات الأفضل في موضوع الدراسة.
2- التحقق من مدى وجود فروق ذات دلالة بين المجموعات فيما يتعلق بالمتغيرات.
3- تحديد المتغيرات التي تساهم بأكبر قدر من الاختلاف بين فئات المتغير التابع.
4- تقسيم المتغير التابع إلى فئات بناء على قيم المتغيرات المستقلة.
5- تقييم دوال التصنيف (معنوية أو غير معنوية).
6- تصنيف مشاهدات جديدة إلى أحد مجموعات المتغير التابع في حالة عدم المعرفة المسبقة بالمجموعات التي تنتمي اليها تلك المفردات.

4- شروط التحليل التمييزي: (بسيوني، 2021)

1. اختلاف متوسطات المجموعات: يجب أن تختلف متوسطات المجموعات لفئات المتغير التابع بشكل واضح.
2. تساوي مصفوفة التباين والتغاير: من الضروري أن تكون مصفوفات التباين والتغاير متساوية بين المجموعات.
3. فصل المجموعات: يجب أن تكون المجموعات منفصلة وقابلة للتحديد بشكل واضح.
4. التوزيع الطبيعي: يجب أن تتبع المتغيرات الكمية والتابعة توزيعاً طبيعياً.
5. العشوائية: يتم اختيار العينة عشوائياً لضمان تمثيل البيانات بشكل مناسب.
6. استقلال المشاهدات: يجب أن تكون المتغيرات المستخدمة مستقلة عن بعضها، حيث يؤدي وجود ارتباط (Multicollinearity) بين المتغيرات إلى صعوبة تفسير النتائج وتحديد المساهمة النسبية لكل متغير.
7. عدم وجود قيم متطرفة: يجب تجنب القيم الشاذة، لأن تحليل التمييز حساس جداً لهذه القيم، وقد تؤدي إلى ابتعاد توزيع البيانات عن التوزيع الطبيعي.

5- معادلة التحليل التمييزي (Discriminant Function):

تكتب الدالة التمييزية على الشكل التالي: (Hair et al, 2019)

$$D = b_1 x_1 + b_2 x_2 + \cdots + b_p x_p$$

D: الدالة التمييزية.

b_p: الأوزان التي يتم تقديرها باسنخدام التحليل التمييزي والتي تستخدم للحصول على تمييز أفضل بين المجتمعين.

x_p: المتغيرات المستقلة وهي مجموعة القياسات التي يتم الحصول عليها.

نستطيع الحصول على القيم المتوقعة لأي فرد من مجموعة من $(n_1 + n_2)$ إذا كانت قيم p معروفة، وللفرد المختار (K) من المجتمع (i)، من خلال الشكل التالي: (الراوي، 2002)

$$D_{ik} = b_1 x_{i1k} + b_2 x_{i2k} + \cdots + b_p x_{ipk}$$

يمكن الحصول على القيمة المتوقعة من خلال هذه الدالة من p من المتغيرات، نستطيع من خلال هذه الدالة تصنيف الفرد إلى المجتمع (من بين مجتمعين أو أكثر) بالاعتماد على قيم $[D_{11}, D_{12}, \dots, D_{1n}]$ إذا كان الفرد من المجتمع الأول، وقيم $[D_{21}, D_{22}, \dots, D_{2n}]$، في هذه الحالة يعطى المجموع الكلي للتباين بالصورة الآتية:

$$T = \sum_{i=1}^{2} \sum_{k=1}^{n_i} (D_{ik} - \overline{D})$$

$$\overline{D} = \frac{1}{n_1 + n_2}\sum_{i=1}^{2}\sum_{k=1}^{n_i} D_{ik} = \frac{1}{n_1 + n_2}(n_1\overline{D}_1 + n_2\overline{D}_2\ )$$

ويمكن تقسيم مجموع المربعات الكلي إلى قسمين:

مجموع المربعات بين المجموعات، ويعطى بالشكل الآتي:

$$B = \sum_{i=1}^{2} n_i\,(D_i - D) = n_1(D_1 - \overline{D}) + n_2(D_2 - \overline{D}) = \frac{n_1.n_2}{n_1 + n_2}(\overline{D}_1 + \overline{D}_2)$$

ومجموع المربعات داخل المجموعات، يعطى بالشكل التالي:

$$W = \sum_{i=1}^{2}\sum_{k=1}^{n_i}(D_{ik} - \overline{D}_i)^2$$

تزداد قيمة B عندما يكون متوسط D في المجموعة الأولى مختلفاً تماماً عن متوسطه في المجموعة الثانية، مما يعكس الفروقات بين المجموعات. يتم حساب الإحصائية W من خلال مجموع المربعات لـ D نسبةً إلى متوسطاتها داخل كل مجموعة، ثم يتم دمجها مع مجموع المربعات داخل المجموعات. يمكن استخدام النسبة B/W كمقياس لتمييز D، حيث تعكس مدى تأثير الاختلافات بين المجموعات مقارنةً بالتغيرات داخل كل مجموعة، وهو الأمر الأساسي الذي يتم التركيز عليه في التحليل.

نفترض أن قيم متوسط المشاهدات للمتغير (j) من القيم المشاهدة من العينة n_i من المجتمع i، على الشكل التالي: $\bar{x}_{ij} = \frac{\sum_{i=1}^{n_i} x_{ijk}}{n_i}$، وبفرض أن قيم فروقات المتوسطات للمتغير (j) من كلا المجموعتين على الشكل التالي:

$$d = \bar{x}_{1j} - \bar{x}_{2j}$$

وبفرض أن:

$$S_{jj'} = \frac{1}{n_1 + n_2 - 2}\sum_{i=1}^{2}\sum_{k=1}^{n_i}\left(\bar{x}_{ijk} - \bar{x}_{ij}\right)\left(x_{ij'k} - x_{ij'}\right)$$

$S_{jj'}$ هو التباين المشترك المدمج للعينات بين المتغيرات j و j' ، حيث يتم تقدير مصفوفة التباين والتباين المشترك:

$$S = \begin{bmatrix} s_{11} & s_{12} & . & . & s_{1p} \\ s_{21} & s_{22} & . & . & s_{2p} \\ . & . & . & . & . \\ . & . & . & . & . \\ s_{p1} & s_{p2} & . & . & s_{pp} \end{bmatrix}$$

يمكن كتابة صيغة عامة لحساب B/W في الصورة التالية:

$$B/W = \frac{\frac{n_1.n_2}{n_1 + n_2}\sum_{j=1}^{p}\sum_{j=1}^{p}\beta_j.\beta_{j'}d_{j.}d_{j'}}{n_1 + n_2 - 2\sum_{j=1}^{p}\sum_{j'=1}^{p}\beta_j.\beta_{j'}S_{jj'}}$$

عندما تكون $p=2$ ، يمكن كتابة الصيغة السابقة بالصورة الآتية:

$$B/W = \frac{\frac{n_1.n_2}{n_1 + n_2}\left({\beta_1}^2 d_1^2 + {\beta_2}^2 d_2^2 + 2\beta_1.\beta_2 d_{1.}d_2\right)}{(n_1 + n_2 - 2)({\beta_1}^2.S_{11} + {\beta_2}^2.S_{22} + 2\beta_1.\beta_2 S_{12}}$$

إن إضافة قيمة ثابتة مثل (C) إلى المعلمات لا تؤثر على (B/w). أي أنه في حال استبدال B_1 $_1$ بدلاً من B_1، و C_2B_2 بدلاً من B_2، ...الخ، وهذا يعني إذا كانت قيم ($_1, b_2, \dots, b_p$) تعظم المقدار (B/w)، فإن مجموع المعلمات ($cb_1, c\ b_2, \dots, cb_p$) هي أيضاً تقوم بتعظيم المقدار (B/w).
حيث خصائص قيم التباين ترتبط ارتباطاً وثيقاً بحجم المعلمات، وهي القيم التي نستخدمها في دالة التمييز عند تصنيف الأفراد إلى مجموعات.
إن مصفوفة المعكوس المرافق إلى مصفوفة S:

$$S^{-1} = \begin{bmatrix} s^{11} & s^{12} & . & . & s^{1p} \\ s^{21} & s^{22} & : & : & s^{2p} \\ . & . & . & . & . \\ s^{p1} & s^{p2} & . & . & s^{pp} \end{bmatrix}$$

وإن القيم ($b_1, b_2, \dots, b_p$) التي تعظم (B/w) يمكن حسابها كالآتي:

$$\left.\begin{array}{llll} b_1 = & s^{11}d_1 + & s^{12}d_2 + & \cdots \quad +s^{1p}d_p \\ b_2 = & s^{21}d_1 + & s^{22}d_2 + & \cdots \quad s^{2p}d_p \\ . & . & . & . \quad . \\ . & . & . & . \quad . \\ b_p = & s^{p1}d_1 + & s^{p2}d_2 & . \quad s^{pp}d_p \end{array}\right\}$$

نرمز للدالة الخطية المعتمدة على $(\grave{b}s)$:

$$D = b_1x_1 + b_2x_2 + \cdots + b_px_p$$

وبطريقة المصفوفات تتم كتابتها بالشكل الآتي:

$$D = (\bar{X}_1 - \bar{X}_2)S^{-1}X$$

6- التحليل التمييزي الخطي (LDA):
يعتمد التحليل التمييزي الخطي على افتراض أن البيانات داخل كل فئة تتبع توزيعات متعددة المتغيرات ذات تباين ثابت (تجانس التباين) بين المجموعات. في هذا النوع، يتم رسم خط أو سطح لفصل بين الفئات بشكل خطي. يستخدم LDA بشكل أساسي عندما تكون الفئات متمايزة بشكل جيد وخصائصها غير معقدة بشكل كبير. يهدف LDA إلى العثور على الخطوط التي توفر أفضل فصل بين المجموعات بناءً على مزيج من الخصائص. (McLachlan,2004)

7- التحليل التمييزي غير الخطي (QDA):
في المقابل، يستخدم التحليل التمييزي غير الخطي عندما لا تنطبق فرضيات التباين الثابت بين الفئات. يسمح QDA باستخدام دوال غير خطية لفصل المجموعات، مما يجعله أكثر مرونة في التعامل مع الحالات التي تكون فيها الفئات أكثر تعقيداً أو تداخلاً. ومع ذلك، يتطلب QDA أعداداً أكبر من المشاهدات لتحديد التباين داخل كل فئة. (Hastie et al, 2009)

8- **اختبار معنوية الدالة المميزة الخطية:** (Hair et al., 2019; Rencher, 2002)
عند التمييز بين مجموعتين، يمكننا أن نختبر الفرضية التي تنص على تساوي متوسطات المجموعات:

$$H_0: \mu_1 = \mu_2 = \mu_3 = \cdots = \mu_k$$

أي جميع المتوسطات متساوية.

$$H_1: \mu_1 \neq \mu_k$$

على الأقل متوسط واحد مختلف.

واختبار معنوية الدالة التمييزية يتوقف على رفض أو قبول فرضية العدم H_0 التي تنص على تساوي متوسطات المجموعات. والاختبار المستخدم للتمييز بين المجموعات هو اختبار فيشر F والذي يعتمد على إحصاءه الاختبار T^2 (*Hoteling*) ، وصيغته الرياضية بالشكل التالي:

$$F = \frac{n_1 + n_2 - p - 1}{(n_1 + n_2 - 2)p} . T^2$$

$$T^2 = \frac{n_1 . n_2}{n_1 + n_2} . D^2$$

$$D^2 = (\bar{X}_1 - \bar{X}_2)' S^{-1} X (\bar{X}_1 - \bar{X}_2)$$

نقوم برفض فرضية العدم عندما تكون قيمة $F_{statistic}$ المحسوبة أكبر من قيمة (F) الجدولية، مما يعني قبول الفرضية البديلة H_1عند مستوى دلالة معين ووفقاً لدرجات الحرية $(n_1 + n_2 - 1)$ يشير ذلك إلى وجود فروق معنوية بين متوسطات المجموعتين، مما يؤكد أن الدالة التمييزية الخطية تتمتع بقدرة عالية على التمييز بينهما.

9- مقياس الاحتمال الصحيح Likelihood Ratio Test:

يستخدم اختبار مقياس الاحتمال الصحيح لتقييم مدى جودة النموذج التمييزي، ويحسب باستخدام نسبة الاحتمال للمجموعات المفترضة مقابل الاحتمال للمجموعات الفعلية. يعطى بالعلاقة: Huberty & (Olejnik, 2006)

$$LR = -2 \sum \ln \left(\frac{L_0}{L_1} \right)$$

حيث:

L_0 هو الاحتمال عندما لا يوجد تأثير للمتغيرات المستقلة.

L_1 هو الاحتمال عند تضمين المتغيرات المستقلة.

كلما كانت القيمة أقل، كان النموذج أكثر قدرة على التمييز بين المجموعات.

10- خطوات التحليل التمييزي:

تتمثل خطوات التحليل التمييزي في النقاط التالية: (العلي، 2019)

1. تحديد نوع التحليل التمييزي المناسب: يتم اختيار النوع الأنسب بناءً على الأهداف الإحصائية المطلوبة، مثل تحديد ما إذا كان التحليل الوصفي أو التنبؤي هو الأنسب، أو ما إذا كان التحليل بسيطاً أم متعدداً، خطياً أم غير خطي، كمياً أم نوعياً، أو لوجستياً.
2. تحديد المتغيرات المستقلة المناسبة: اختيار المتغيرات المستقلة الضرورية للتحليل التمييزي الذي تم اختياره، والتي تساهم في تحقيق أهداف البحث، ومن ثم جمع البيانات المتعلقة بها.
3. تحديد عدد المجموعات في المجتمع: تحديد عدد المجموعات التي ستستخدم في التحليل، وتحديد ما إذا كانت هناك مجموعات ثنائية أو متعددة، ثم سحب العينات الطبقية منها وجمع البيانات اللازمة.
4. اختبار البيانات: التأكد من أن البيانات تتوافق مع الافتراضات والشروط اللازمة لإجراء التحليل التمييزي بشكل صحيح.
5. إجراء التحليل التمييزي: تطبيق خطوات التحليل التمييزي العملية للوصول إلى النتائج الإحصائية المطلوبة.
6. تفسير النتائج: تحليل النتائج المستخلصة والعمل على استثمارها لتحقيق الفائدة المرجوة من التحليل.

المبحث الثاني: الجانب التطبيقي

تمهيد:

تم الاعتماد في هذه الدراسة على بيانات اقتصادية مستخلصة من قاعدة بيانات البنك الدولي لعام 2023. شملت البيانات متغير معدل النمو الاقتصادي (نمو إجمالي الناتج المحلي %)، إجمالي تكوين رأس المال (% من إجمالي الناتج المحلي)، نصيب الفرد من إجمالي الناتج المحلي (بالأسعار الجارية للدولار الأمريكي).

تم تصنيف الدول إلى مجموعتين: دول ذات نمو اقتصادي مرتفع ودول ذات نمو اقتصادي منخفض، وذلك بناءً على متوسط معدل النمو الاقتصادي العالمي.

تم تطبيق التحليل التمييزي الخطي باستخدام برنامج SPSS لتقييم قدرة المتغيرات الاقتصادية المختارة على التمييز بين المجموعتين. وتم اختبار معنوية النموذج باستخدام اختبار Wilks' Lambda لتحديد مدى فعالية التحليل التمييزي في تصنيف الدول بدقة.

الجدول (1): تصنيف الدول بناءً على معدل النمو الاقتصادي.

البلد	إجمالي تكوين رأس المال (% من إجمالي الناتج المحلي)	نصيب الفرد من إجمالي الناتج المحلي (بالأسعار الجارية للدولار الأمريكي)	النمو الاقتصادي %	المجموعة
الصين	42.07895156	12614.06	5.36	مرتفع
إندونيسيا	30.5197744	4876.31	0.61	مرتفع
البرازيل	16.08598735	10294.87	2.50	مرتفع
فيتنام	31.96175139	4282.09	4.34	مرتفع
الهند	33.31986126	2480.79	7.20	مرتفع
تركيا	29.96416793	13105.66	4.68	مرتفع
الإمارات العربية المتحدة	27.07865632	49040.69	3.62	مرتفع
الجزائر	37.72494621	5364.03	4.10	مرتفع
ليبيا	16.48035514	6172.81	10.16	مرتفع
مصر	12.87539257	3457.46	3.76	مرتفع
السعودية	29.34212012	32093.96	-0.75	منخفض
الأرجنتين	19.37153901	14187.48	3.64	منخفض
جنوب أفريقيا	15.53809338	6022.54	-0.63	منخفض
ألمانيا	21.67254005	54343.23	0.35	منخفض
إيطاليا	22.75497093	39003.32	0.73	منخفض
فرنسا	23.15336445	44690.93	0.61	منخفض
اليونان	16.73123276	23400.73	2.64	منخفض
النرويغ	25.87601678	87925.09	-0.66	منخفض
كندا	24.0198953	53431.19	-1.68	منخفض

المصدر: بيانات البنك الدولي

https://data.albankaldawli.org/indicator/NY.GDP.MKTP.KD.ZG

المتغيرات المستخدمة في الدراسة:

المتغير المعتمد الذي يمثل المجموعة: نمو اقتصادي منخفض (0)، نمو اقتصادي مرتفع (1)

المتغيرات المستقلة:

إجمالي تكوين رأس المال (% من إجمالي الناتج المحلي)

نصيب الفرد من إجمالي الناتج المحلي (بالأسعار الجارية للدولار الأمريكي)

النمو الاقتصادي %

الجدول (2): المتوسطات الحسابية للمتغيرات المستقلة.

المتوسطات / المجموعات	إجمالي تكوين رأس المال (% من إجمالي الناتج المحلي)	النمو الاقتصادي	نصيب الفرد من إجمالي الناتج المحلي (بالأسعار الجارية للدولار الأمريكي)
منخفض	22.05	0.47	39455.39
مرتفع	27.81	4.63	11168.88
المتوسط العام	25.08	2.66	24567.75

المصدر: من إعداد الباحثين بالاعتماد على SPSS20

اختبار التوزيع الطبيعي:

تم اختبار التوزيع الطبيعي للمتغيرات المستقلة باستخدام اختبار كولموغروف-سميرنوف. أظهرت نتائج الاختبار أن قيم p value كانت غير معنوية لجميع المتغيرات (أكبر من 0.05)، مما يشير إلى أن المتغيرات تتبع التوزيع الطبيعي.

الجدول (3): نتائج اختبار التوزيع الطبيعي.

	إجمالي تكوين رأس المال (% من إجمالي الناتج المحلي)	النمو الاقتصادي	نصيب الفرد من إجمالي الناتج المحلي (بالأسعار الجارية للدولار الأمريكي)
Kolmogorov-Smirnov Z	.497	.692	1.073
Asymp. Sig. (2-tailed)	.966	.725	.200

المصدر: من إعداد الباحثين بالاعتماد على SPSS20

اختبار تساوي مصفوفات التباين والتغاير:

تم اختبار مساواة التباين المشترك بين المجموعات باستخدام اختبار Box's M. أظهرت النتائج أن القيمة p كانت 0.094، وهي غير معنوية (أكبر من 0.05)، مما يشير إلى قبول فرضية مساواة التباين المشترك بين المجموعات. بناءً على ذلك، يمكن متابعة التحليل التمييزي.

الجدول (4): اختبار Box's M.

Box's M		**13.459**
F	Approx.	1.806
	df1	6
	df2	2018.597
	Sig.	.094

المصدر: من إعداد الباحثين بالاعتماد على SPSS20

اختبار معنوية الدالة التمييزية:

تم اختبار فرضية تساوي متوسطات المجموعات المعرفة بالصيغة التالية:

$H_0: \mu_1 = \mu_2$ متوسط المجموعة الأولى يساوي متوسط المجموعة الثانية

$H_1: \mu_1 \neq \mu_2$ متوسط المجموعة الأولى لا يساوي متوسط المجموعة الثانية

بلغت قيمة Wilks' Lambda 0.374، كما بلغت قيمة (Sig. = 0.002) وهي أصغر من 0.05 وبالتالي نرفض فرضية العدم H_0 والتي تنص على عدم وجود فروق معنوية بين متوسط المجموعة الأولى ومتوسط المجموعة الثانية،وهذا يد على معنوية الفرق، أي أن هناك تمييز بين المجموعات، هذه النتائج تدل على أن التحليل التمييزي يتمتع بقدرة جيدة على التمييز بين المجموعتين حسب المتغيرات المستقلة.

الجدول (5): نتائج Wilks' Lambda

Test of Function(s)	Wilks' Lambda	Chi-square	df	Sig.
1	.374	15.264	3	.002

المصدر: من إعداد الباحثين بالاعتماد على SPSS20

اختبار معنوية المتغيرات في الدالة التمييزية:

لمعرفة أهمية وتأثير كل متغير من المتغيرات المستقلة، تم اختبار معنوية هذه المتغيرات من خلال تحليل التباين الأحادي.

يوضح الجدول (6) أن المتغير (النمو الاقتصادي)، والمتغير (نصيب الفرد من إجمالي الناتج المحلي) يتمتعان بمعنوية عالية، بينما متغير (إجمالي تكوين رأس المال) ليس له تأثير معنوي في عملية التصنيف.

الجدول (6): تحليل معنوية المتغيرات المستقلة.

variables	Wilks' Lambda	F	df1	df2	Sig.
إجمالي تكوين رأس المال (% من إجمالي الناتج المحلي)	.865	2.664	1	17	.121
النمو الاقتصادي	.506	16.593	1	17	.001
نصيب الفرد من إجمالي الناتج المحلي (بالأسعار الجارية للدولار الأمريكي)	.635	9.787	1	17	.006

المصدر: من إعداد الباحثين بالاعتماد على SPSS20

يُظهر تحليل Structure Matrix أن المتغيرات الأكثر تأثيرًا في تصنيف الدول وفقًا للنمو الاقتصادي هي نصيب الفرد من إجمالي الناتج المحلي والنمو الاقتصادي. حيث حقق نصيب الفرد من إجمالي الناتج المحلي أعلى ارتباط سلبي مع الدالة التمييزية (-0.586)، مما يشير إلى دوره القوي في التمييز بين المجموعتين. في المقابل، جاء تأثير النمو الاقتصادي إيجابيًا لكنه أقل قوة (0.306)، مما يعني أنه يساهم في التصنيف ولكن بدرجة أقل مقارنةً بنصيب الفرد من إجمالي الناتج المحلي. أما إجمالي تكوين رأس المال، فقد لم يظهر في Structure Matrix، مما يدل على ضعف تأثيره في عملية التصنيف. كما أن هذا المتغير كان غير معنوي في الجدول السابق، مما يفسر عدم ظهوره في المصفوفة واستخدامه المحدود في التمييز بين المجموعات.

الجدول (7): العلاقة بين المتغيرات والدالة التمييزية Structure Matrix.

	Function1
النمو الاقتصادي	.306
نصيب الفرد من إجمالي الناتج المحلي (بالأسعار الجارية للدولار الأمريكي)	-.586-

المصدر: من إعداد الباحثين بالاعتماد على SPSS20

مقياس الاحتمال الصحيح:

يوضح الجدول (8) فعالية التحليل التمييزي في تصنيف المجموعتين، تمكن النموذج من تصنيف 82.4% من الحالات الأصلية بشكل صحيح.

بالنسبة للمجموعة "منخفض" بلغت دقة التصنيف 88.9%

بالنسبة للمجموعة "مرتفع" بلغت دقة التصنيف 80%

تشير هذه النتائج إلى ان التحليل التمييزي نجح بشكل كبير في تصنيف بعض الدول إلى مجموعتين حسب معدل نمو اقتصادي مرتفع بناءً على عدة متغيرات مستقلة، مما يدل على قوة النموذج وقدرته العالية على التمييز بشكل دقيق.

الجدول (8): نتائج التصنيف Classification Results[a]

المجموعة			Predicted Group Membership		Total
			منخفض	مرتفع	
Original	Count	منخفض	8	1	9
		مرتفع	2	8	10
	%	منخفض	88.9	11.1	100.0
		مرتفع	20.0	80.0	100.0

a. 82.4% of original grouped cases correctly classified.

المصدر: من إعداد الباحثين بالاعتماد على SPSS20

النتائج:

1- أظهرت النتائج أن التحليل التمييزي كان قادراً على تصنيف 82.4% من الحالات الأصلية بشكل صحيح، مما يعكس قوة النموذج في التنبؤ بدقة عالية وتصنيف الدول وفقاً لمعدلات نموها الاقتصادي. تشير هذه النتيجة إلى أن النموذج المستخدم يمتلك مستوى عالٍ من الكفاءة والموثوقية، مما يجعله أداة تحليلية مناسبة لاستخدامها في الدراسات الاقتصادية التي تتطلب تصنيف الدول بناءً على مؤشرات اقتصادية محددة.

2- أظهرت نتائج التحليل التمييزي قدرة متميزة على التفرقة بين الدول ذات النمو الاقتصادي المرتفع والدول ذات النمو الاقتصادي المنخفض بناءً على معدل النمو الاقتصادي. حيث بلغت قيمة الدلالة الإحصائية لقيمة Wilks' Lambda (Sig = 0.002)، وهي أقل من مستوى الدلالة 0.05، مما يعني رفض الفرضية الصفرية (H_0) التي تفترض عدم وجود فروق معنوية بين المجموعتين. ويدل ذلك على أن هناك اختلافاً جوهرياً بينهما، مما يؤكد فعالية التحليل التمييزي في تصنيف الدول وفقاً للمتغيرات المستخدمة.

3- تبين أن النمو الاقتصادي ونصيب الفرد من إجمالي الناتج المحلي يتمتعان بمستوى معنوية مرتفع، مما يشير إلى أهميتهما الكبيرة في التمييز بين الدول. على العكس من ذلك، لم يكن لمتغير إجمالي

تكوين رأس المال تأثير معنوي على عملية التصنيف، مما يعني أن هذين المتغيرين لا يسهمان بشكل جوهري في تحديد الفرق بين المجموعتين.

4- التحليل التمييزي يوفر أداة فعالة لفهم الفروق الاقتصادية بين الدول، حيث يمكن الاعتماد عليه في تصنيف وتصنيف الدول وفقاً لمؤشرات النمو الاقتصادي، مما يسهم في دعم عملية صنع القرار الاقتصادي والسياسات التنموية.

التوصيات:

1- تعزيز الاعتماد على التحليل التمييزي في الدراسات الاقتصادية، نظراً لفعالية النموذج في تصنيف الدول وفقاً لمعدلات النمو الاقتصادي، مما يساهم في صياغة سياسات تنموية أكثر دقة.

2- إجراء دراسات مقارنة بين عدة طرق إحصائية حيث يمكن للباحثين في المستقبل مقارنة نتائج التحليل التمييزي مع تقنيات أخرى مثل تحليل الانحدار أو أساليب التقدير غير الخطية لفهم ما إذا كانت الطرق الأخرى تقدم تحليلات أفضل أو أكثر دقة.

3- نظراً لدقة النموذج في التصنيف (82.4%)، يمكن الاستفادة منه في التنبؤ بأداء الدول اقتصادياً، مما يساعد في توجيه السياسات الاقتصادية والتخطيط الاستراتيجي للدول النامية.

4- يوصى بمقارنة نتائج التحليل التمييزي مع نماذج أخرى مثل الانحدار اللوجستي أو الشبكات العصبية الاصطناعية، وذلك لتحسين دقة التصنيف وفهم العلاقات الاقتصادية بشكل أكثر تعمقاً.

5- يوصى بتوسيع نطاق العينة في الدراسات المستقبلية لتشمل مزيداً من الدول والاقتصادات، بما في ذلك الدول ذات التنوع الكبير في الأداء الاقتصادي. هذا سيسمح بتحليل أكثر شمولاً ويزيد من دقة النموذج.

المراجع

المراجع العربية:

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

2- الراوي، عمر فوزي صالح بدر. (2004). **استخدام الدالة التمييزية في السيطرة النوعية مع تطبيق على ولادات الاطفال الخدج**. رسالة مقدمة لنيل درجة الماجستير علوم في الإحصاء، جامعة الموصل، الموصل، العراق.

3- الزاوي، إيمان موسى فرج. (2023). **فاعلية نموذج التحليل التمييزي في تصنيف ودراسة العوامل المؤثرة في تكيف الطلبة الوافدون بجامعة إجدابيا**. *مجلة رماح للبحوث والدراسات*، (78)، 27-48.

4- الشمراني ، محمد (٢٠٠٩) **دراسة مقارنة بين التحليل التمييزي وتحليل التباين المتعدد في تحليل البيانات متعددة المتغيرات**. أطروحة دكتوراه غير منشورة، جامعة أم القرى، مكة المكرمة، المملكة العربية السعودية.

5- الشمراني، محمد موسى. (2020). **توظيف أسلوبي التحليل العنقودي والتحليل التمييزي في تصنيف البيانات وبناء الدوال التمييزية**.مجلة كلية التربية - جامعة الأزهر، 39(186)،11-39

6- صالح، عائدة هادي. (2008). **استخدام التحليل المميز لتشخيص بعض أمراض العيون**.مجلة الإدارة والاقتصاد، 67، 264 - 287

7- عاشور، ضياء الدين عبد الرازق عبد الغني؛ مهدي، مرفت مهدي رمضان؛ هنداوي، محمد جودة خليل. (٢٠٢٤). **استخدام التحليل الاحصائي متعدد المتغيرات الدراسة الهيكل البنائي الاقتصادي والاجتماعي للممولين دراسة تطبيقية**، المجلة العلمية للدراسات والبحوث المالية والتجارية، كلية التجارة، جامعة دمياط، ٥(١)، ٤، ٣٥١-٣٨١

8- العلي، إبراهيم محمد. (2019). **الأسس الرياضية للتحليل التمييزي الخطي والنوعي الأسس الرياضية للتحليل اللوجستي والتحليل العنقودي**. سورية: جامعة تشرين.

9- النعيمي، عز الدين عبد الله. (2019). **فاعلية أسلوب التحليل التمييزي في التنبؤ بالتعثر الدراسي من خلال اختبار القدرات العامة والاختبار التحصيلي في بعض الجامعات السعودية**. قسم علم النفس كلية العلوم الاجتماعية، جامعة الإمام محمد بن سعود الإسلامية.
10- نده، كنده عبد الحميد. (2019). **استخدام التحليل التمييزي لتصنيف المستوى المعيشي للأسر في محافظة اللاذقية**. المجلة الجزائرية للأبحاث الاقتصادية والمالية. 2(2)، 11-39.
11- البنك الدولي. (2023). معدل نمو الناتج المحلي الإجمالي (%): بيانات عام 2023. تم الاسترجاع في 15 يناير 2025 من:
https://data.albankaldawli.org/indicator/NY.GDP.MKTP.KD.ZG

المراجع الأجنبية:

1. Erimafa J.T., Iduseri A. and Edokpa I.W. (2009). **Application of discriminant analysis to predict the class of degree for graduating students in a university system**, International Journal of Physical Sciences, 4(1), 16-21.
2. Johnson, R.A. and Wichern, D.W. (2002**). Applied Multivariate Statistical Analysis**. Prentice Hall, New Jersey.
3. Hastie, T., Tibshirani, R., & Friedman, J. (2009). **The Elements of Statistical Learning: Data Mining, Inference, and Prediction** (2nd ed.). Springer.
4. Huberty, C. J., & Olejnik, S. (2006). **Applied MANOVA and Discriminant Analysis** (2nd ed.). Wiley.
5. Hardigan, P., Lai, L., Arneson, D. & Robeson, A. (2001). **Significance of Academic Merit, Test Scores, Interviews and the Admissions Process: A Case Study**. American Journal of Pharmaceutical Education, 65, 40-44.
6. Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). **Multivariate Data Analysis** (8th ed.). Cengage Learning.
7. McLachlan, G. J. (2004). **Discriminant Analysis and Statistical Pattern Recognition**. Wiley-Interscience.
8. Rencher, A. C. (2002). **Methods of multivariate analysis**. (2nd ed.). Wiley.
9. Stella, O. (2019). **Discriminant Analysis: An Analysis of Its Predictship Function.Journal of Education und Practice**, 10(5), 50-57