استخدام الشبكات العصبونية لدعم قرار تصنيف الكفالات المالية (بالتطبيق على الجمعيات الخيرية في مدينة حلب)

Using Artificial Neural Networks to Support Financial Sponsorship Classification Decision (Applied on Charities in Aleppo)

محمد لؤي عبد الرزاق دهان
طالب (دكتوراه) / جامعة حلب/ سورية

ياسر عبد الله الموسى
أستاذ دكتور/ جامعة حلب/ سورية

Mohamed Loay Abdulrazak Dahhan
PhD.Student/ Aleppo University/ Syria
mloaydahhan@alepuniv.edu.sy

Yasser Al-Mousa
Professor/ Aleppo University/ Syria
ymoussa@alepuniv.edu.sy

تاريخ الاستلام: 6/ 11/ 2019م، **تاريخ القبول:** 23/ 5/ 2020م.
E-ISSN: 2410-3349
P-ISSN: 2313-7592

Received: 6/ 11/ 2019, **Accepted:** 23/ 5/ 2020.
DOI: 10.33977/1760-005-013-010
https: //journals.qou.edu/index.php/eqtsadia

ملخص:

يهدف هذا البحث إلى بناء نموذج باستخدام الشبكة العصبونية الصناعية لمساعدة متخذي القرار على اتخاذ قرار تصنيف الكفالات المالية في الجمعيات الخيرية بحلب من خلال استخدام المنهج الوصفي التحليلي ومن ثم بناء شبكة عصبونية متعددة الطبقات ANN تعتمد على قاعدة بيانات متعددة الأبعاد Multivariat مختلفة النوع والطبيعة، ومقارنة النتيجة مع تحليل التمايز Discriminant Analysis كأحد أبرز أساليب التصنيف Classification التقليدية، طبقت الدراسة على عدد من الجمعيات الخيرية في مدينة حلب والمنتسبة إلى اتحاد الجمعيات الخيرية للأعوام (2016 – 2019) والتي تحوي على قسم الكفالات المالية، وتوصلت الدراسة إلى أهمية استخدام الشبكة العصبونية في اتخاذ قرار تصنيف الكفالات المالية في الجمعيات الخيرية حيث كانت نسبة التصنيف الصحيحة فيها أعلى وبشكل واضح من تحليل التمايز، وبينت أهمية متغيري العمر وعدد المكفولين لاتخاذ قرار الكفالة، مما يؤدي إلى اتخاذ قرار سليم بنسبة أعلى وبالوقت المناسب.

الكلمات المفتاحية: شبكات عصبونية، تحليل التمايز، تصنيف، جمعيات خيرية.

Abstract:

This paper aims to build a model using the artificial neural network (ANN) to help decision makers to decide the classification of financial guarantees in charities in Aleppo. By using the descriptive analytical method and building a multi - layered neural network based on multivariate database of different kind and nature, and comparing the result with Discriminant Analysis as one of the most Classification methods. This study was applied to a number of charities in Aleppo affiliated to the Union of charities for the years 2016 - 2019, which contain the section of financial guarantees. The study found that the importance of using the neural network in making decision to classify financial guarantees in charities, where the correct rating in classification which was clearly higher than the discriminant analysis. In addition, the study showed the importance of the variables of age and the number of sponsored to make a decision, which leads to making decisions with the highest rate and in a timely manner.

***Keywords**: Neural Network, Discriminant Analysis, Classification, Charities.*

مقدمة البحث:

إن عمل معظم الجمعيات الخيرية يتمركز بشكل أساسي في جمع التمويل اللازم من قبل المتبرعين، وتوزيع ما تم تجميعه ليحقق أفضل عائد ممكن على كل من الفرد والمجتمع، وبالتالي فإن الإدارة الحكيمة لاتخاذ القرارات المتعلقة بإعادة توزيع الأموال من أهم الأنشطة الواجب الاعتناء بها.وعليه فإن عملية تطوير أنظمة دعم القرارات DSSs أصبح ضرورة ملحة على أصحاب القرار، ولاسيما في ظل الظروف التي مرت على مجتمع البحث جعلت من عملية اتخاذ القرار أمرا غاية في التعقيد.

إن خصوصية مثل هذا النوع من القرارات يستوجب الإلمام بكافة البيانات الممكن أن تساهم باتخاذ القرار بعيداً عن التحيز والرأي الشخصي، ولما كانت الطرق الإحصائية التقليدية تستوجب شروطاً معينة لبناء نماذج تساعد في اتخاذ قرارات من هذا النوع، كان لابد من البحث في الطرق الحديثة عن أداوت يمكن أن تنافس الطرائق السابقة.كما أن الأبعاد التكنولوجية الجديدة وتطور تقنيات الحوسبة والذكاء الصنعي جعلت من عملية اتخاذ القرار أمراً أكثر سهولة ويسراً، وأصبح تطبيق هذه الأبعاد واسعاً أكثر وبما يخدم أصحاب القرار.ومن هنا تطورت الشبكات العصبونية الصناعية لتساعد في عملية اتخاذ القرار بما في ذلك القرارات التي تستوجب التنبؤ أو التصنيف.

مشكلة البحث:

يصادف متخذ القرار أثناء اتخاذ قرار لتصنيف كفالة شخص جديد عدة نقاط أبرزها معالجة البيانات ذات الأبعاد المتعددة Multivariat Dataset الخاصة بطالب الكفالة الجديد والأشخاص المكفولين مسبقاً قبل اتخاذ القرار، وعندما يقوم بإدخال كامل أبعاد البيانات للمعالجة فإن ذلك يؤدي إلى زيادة درجة تعقيد عملية اتخاذ القرار المطبق وبالتالي ضياع في الوقت والجهد، وقد يؤدي في بعض الأحيان إلى اتخاذ قرارات غير سليمة، وبسبب تنوع البيانات ذات الأبعاد المتعددة واختلاف طبيعتها يصعب تطبيق التحليلات الإحصائية التقليدية للتصنيف.ومن هنا تبرز المشكلة في معالجة قاعدة البيانات ذات الأبعاد المتعددة والتي تعطي بالنتيجة قرارا للتصنيف غير مقبول وبتقييم ضعيف، والبحث عن أسلوب آخر يقبل البيانات باختلاف طبيعتها ونوعها ويعطي نتيجة مقبولة للتنبؤ في التصنيف مما ينعكس بشكل إيجابي في عملية اتخاذ القرار.وهذا ما يقودنا إلى صياغة مشكلة البحث وفق السؤال التالي:

هل تعطي استخدام الشبكة العصبونية الصناعية متعددة الطبقات نتائج دقيقة في التنبؤ بقرار التصنيف؟

والذي يتفرع عنه السؤالان التاليان:

- ما هي المتغيرات الأكثر تأثيراً في اتخاذ قرار تصنيف الكفالات المالية؟
- ما مدى تطبيق الشبكة العصبونية الصناعية متعددة الطبقات بالمقارنة مع تحليل التمايز في التنبؤ بتصنيف الكفالات المالية؟

بيانات متعددة الأبعاد مقارنة بتحليل التمايز كأحد الطرق الإحصائية المستخدمة في التصنيف، وتحديد المتغيرات ذات الأثر الأكبر في التصنيف والتي تساهم في صنع القرار الخاص بالكفالات المالية في الجمعيات الخيرية.

منهجية البحث:

اعتمد الباحث على المنهج الوصفي التحليلي، ومن ثم بناء شبكة عصبونية متعددة الطبقات وتدريبها للتنبؤ بقرار تصنيف في قاعدة بيانات ذات أبعاد متعددة، حيث سيتم تنفيذ الشبكة العصبونية ومن ثم تطبيق تحليل التمايز على ذات البيانات ومقارنتهما من خلال اعتماد نسبة التصنيف الصحيح كأسلوب لتقييم النتيجة.

حدود البحث:

- الحدود الزمانية: تغطي الفترة الواقعة ما بين 10 – 8 – 2016 و 23 – 10 – 2019.
- الحدود المكانية: أقسام الكفالات المالية في عدد من الجمعيات الخيرية في مدينة حلب – سورية.

مجتمع وعينة البحث:

يشمل مجتمع البحث جميع العائلات المقدمة لطلب الكفالة المالية للجمعيات الخيرية في مدينة حلب والمنتسبة إلى اتحاد الجمعيات الخيرية للأعوام (2016 – 2019)، تم اختيار العائلات المقدمة لطلب الكفالة المالية بشكل عشوائي من قبل 4 جمعيات (وهي الجمعيات التي تحوي على أنظمة معلومات خاصة بقسم الكفالات المالية)، أما بالنسبة للعائلات في الجمعيات الأخرى فيتعذر جمع المعلومات عنها بسبب صعوبة توفر المعلومات اللازمة لديهم، وبالنسبة لحجم العينة يجب أن تحوي على عدد من المفردات يزيد عن عدد الأوزان الكلي في الشبكة من خمسة إلى عشرة مرات، واقترح كل من Baum و Haussler علاقة لتحديد الحدّ الأدنى لعدد المفردات مرتبطاً بنسبة الثقة المطلوبة في العينة حسب (Mehrotra, et al., 1997, P.86 - 88)، هي:

$$p = \frac{|W|}{(1-\alpha)}$$

- إذ إنّ: P عدد مفردات العينة، W عدد الأوزان الكلية في الشبكة، α نسبة الثقة المطلوبة في العينة، والشرط الكافي لضمان الأداء المطلوب يعطى بالعلاقة:

$$p = \frac{|W|}{(1-\alpha)} \; log \frac{n}{1-\alpha}$$

- إذ إنّ n تمثّل عدد العقد في الشبكة، وعليه فإن الحد الأدنى لحجم العينة اللازم عند مستوى ثقة 95% هو:

$$p = \frac{|122|}{(1-0.95)} \; log \frac{20}{1-0.95}$$

$$p = \frac{122}{0.05} * \log(400) = 2440 * 2.60 = \mathbf{6344}$$

الدراسات السابقة:

دراسة (سليمان، 2015) بعنوان: «المقارنة بين التحليل التمييزي والنموذج اللوجستي الثنائي ونماذج الشبكات العصبية في تصنيف المشاهدات: بالتطبيق على دراسة العوامل المؤثرة على كفاية دخل الأسرة»:

تهدف هذه الدراسة إلى التعرف على الأساليب التي تعالج النماذج ذات المتغيرات التابعة النوعية والمفاضلة بينها، وتسليط الضوء على استخدام الشبكات العصبية الاصطناعية في التطبيقات الإحصائية خاصة في التصنيف والتنبؤ، والتعرف على أهم العوامل التي تؤثر على كفاية دخل الأسرة.اعتمدت الدراسة على المنهج الوصفي والتحليلي الاستقرائي لتوصيف المتغيرات وبناء النماذج. وتم تطبيقها في ولاية الخرطوم بين عامي 2012 و 2015.من أهم نتائج الدراسة أن النماذج المقترحة أعطت نتائج متطابقة من حيث معنوية تأثير وأهمية المتغيرات المستقلة الداخلة في التحليل. كما نجد أن طريقة الشبكات العصبية الاصطناعية قد أعطت نسبة تصنيف أفضل من النموذج اللوجستي وطريقة الدالة التمييزية، تم استخدام معيار نسبة المشاهدات المصنفة خطأ (نسبة التصنيف الخاطئ) كمعيار للمقارنة.من أهم توصيات الدراسة الاستفادة من الأساليب الإحصائية المتقدمة (أساليب التحليل متعدد المتغيرات) مثل الدالة التمييزية، ونماذج التصنيف الحديثة مثل نماذج الشبكات العصبية، للفصل أو التمييز بين مجموعتين أو أكثر، في جميع مجالات المعرفة.

دراسة (العدوي، 2017) بعنوان «دراسة نماذج الانحدار للمتغيرات النوعية»، هدفت هذه الدراسة إلى دراسة ومقارنة نماذج الانحدار المستخدمة في تصنيف المتغيرات النوعية وهي التحليل التمييزي Discriminant Analysis وتحليل الانحدار اللوجستي Regression Analysis، تم تطبيق الدراسة على عينتين الأولى لمرضى السكري والثانية لمرضى السرطان، وتوصلت الدراسة إلى عدد من النتائج أهمها أن الاختلاف بين الأسلوبين يكمن في اختلاف شروط تطبيق كل منهما، وأن أسلوب الانحدار اللوجستي يصنف المفردة وفق احتمال معين أما التمييزي فيقوم مباشرة بتصنيف المشاهدة، اعتمد على معامل التحديد R^2 في المقارنة بين الأسلوبين.أوصى الباحث باشراك خوارزميات الذكاء الصنعي المستخدمة في التصنيف ومقارنتها مع الأسلوبين السابقين.

دراسة (Barzanti, et al.– 2016) بعنوان "نظام دعم القرار للمنظمات غير الربحية":

تهدف الدراسة إلى تطوير نموذج للجمعيات متوسطة الحجم غير الربحية، والذي يأخذ بعين الاعتبار خصوصيات هذا النوع من الجمعيات غير الحكومية من أجل تحسين عملية جمع الأموال، اعتمدت الدراسة بشكل أساسي على خوارزمية تصنيف من خلال أداة تم تصميمها، تقوم الأداة بتحليل البيانات الضخمة ومن ثم يتم تصنيف المتبرع بهدف معرفة فيما إذا تم التوصل للهدف اللازم أم لا، تم تطوير الأداة كنظام دعم قرار فعال وتم تطبيقها على قاعدة بيانات ضخمة افتراضية تحاكي آلية عمل هذه الجمعيات، وتوصلت إلى نتائج قريبة بشكل كبير من الواقع من خلال تحليل النتائج المحققة مع النتائج الفعلية.

دراسة (Khemakhem & Boujelbène, 2015) بعنوان "التنبؤ بمخاطر الإقراض: دراسة مقارنة بين التحليل التمييزي وطريقة الشبكات العصبونية":

هدفت هذه الدراسة إلى استكشاف طرق عملية جديدة تعتمد على الشبكات العصبونية لمساعدة متخذ القرار في المصرف للتنبؤ بمخاطر عدم الدفع للشركات التي تطلب القرض، وتمت مقارنة النتائج مع التحليل التمييزي، طبقت هذه الدراسة على 86 شركة زبائن تونسية من قطاعات مختلفة على مدى ثلاث سنوات 2005 إلى 2007. اعتمدت الدراسة على بناء النماذج من خلال المنهج الوصفي التحليلي، حيث تم اختيار المتغيرات اللازمة ومن ثم تقسيم العينة إلى مجموعتين الأولى للتدريب والثانية للاختبار، وتم الاعتماد على نسبة التصنيف الصحيحة للمقارنة بين كلا الأسلوبين، وتوصلت الدراسة إلى أن الشبكة العصبونية مضبوطة بشكل أفضل من التحليل التمييزي في التنبؤ والتصنيف.

دراسة (Leela & Loshma, 2012) بعنوان "التصنيف والتنبؤ في إدارة علاقات العملاء باستخدام نموذج الشبكة العصبونية متعددة الطبقات ذات التغذية العكسية":

ناقشت هذه الدراسة أهمية CRM في المنظمات، وبينت أهمية التصنيف والتنبؤ لمساعدة المنظمة على دراسة وتحليل والتنبؤ بأنماط العملاء، اعتمدت على تقنيات التنقيب في البيانات للتصنيف ومن ثم قامت بتطوير نموذج شبكة عصبونية لمعالجة البيانات بشكل أكثر كفاءة وفعالية، وبوقت تنفيذ أقل.توصلت الدراسة إلى أن الشبكات العصبونية متعددة الطبقات ذات التغذية العكسية تحسن عملية المعالجة لتنفيذ CRM وبوقت أكثر فعالية وكفاءة.

وتتميز الدراسة الحالية باتفاقها مع كل من دراسة (سليمان، 2015) و (العدوي، 2017) و (Khemakhem & Boujelbène, 2015) في استخدام التحليل التمييزي كأبرز الطرق الإحصائية المستخدمة في التصنيف للمقارنة مع نموذج الشبكة العصبونية المقدم في هذه الدراسة.وتتفق مع دراسة (Barzanti, et al.– 2016) من حيث أهمية تطبيق نظم دعم قرارات ونظم معلومات تعتمد على تقنيات الحوسبة المتطورة لتساهم في اتخاذ القرار لدى الجمعيات الخيرية، وتتفق مع دراسة (Leela & Loshma, 2012) بضرورة تطبيق تقنيات التصنيف في الجمعيات الخيرية للسعي للوصول إلى نظام إدارة علاقة مع العميل والمتمثل في المستفيدين من الكفالة المالية.

وبالتالي فإن ما يميز الدراسة الحالية هو استخدام الشبكات العصبونية الصناعية للتنبؤ بقرار تصنيف الكفالات المالية في قاعدة بيانات ذات أبعاد متعددة، وتحديد المتغيرات الأكثر أهمية في عملية اتخاذ قرار التصنيف ومقارنتها مع تحليل التمايز كأحد الأساليب الإحصائية لتبيان جدواها في أقسام الكفالات المالية في الجمعيات الخيرية، للوصول أخيرا إلى نظام إدارة علاقة مع العملاء يكون فعال وداعم للقرارات التي تتخذها إدارة هذه الأقسام.

الإطار النظري للبحث:

◄ **ففهوم الشبكات العصبونية الصناعية Artificial Neural Networks (ANNs):**

تعمل الشبكات العصبونية الصناعية من خلال إنشاء اتصالات Connections بين عناصر المعالجة العصبونية الرياضية Mathematical PEs، حيث يتم ترميز المعرفة داخل الشبكة من خلال الأوزان Weights والتي تمثّل قوّة الاتصال بين العناصر العصبونية، وإنشاء طبقات Layers تجمع هذه العناصر لتعمل مع بعضها على التوازي.

يمكن تعريف الشبكة العصبونية الصناعية على أنها نموذج يُحاكي عمل الشبكة العصبونية الطبيعية باستخدام تقنيات حسابية إنجاز مهمّة معينة، معتمدة على المعالجة المتوازية الموزعة Parallel Processing باستخدام عنصر المعالجة العصبونية Process Element (PE) من خلال تخزين المعرفة والقدرة على التدريب بهدف استخدامها لاحقاً لإنجاز هذه المهمة.

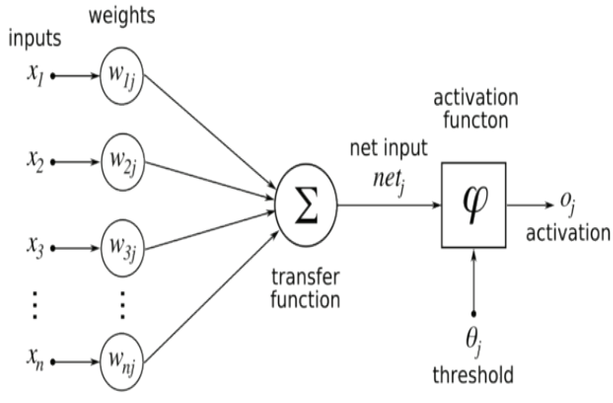
ومن ثم تتعلم الشبكة العصبونية من خلال تقدير عدد العقد وضبط أوزان الاتصالات بالاعتماد على مجموعة التدريب Training Data، وتتعلم الشبكة إما بطريقة التعليم المراقب Supervised Learning وفيها تحوي مجموعة التدريب على أزواج من الأنماط المدخلة والمستهدفة Input - Output Pairs، تحاول الشبكة العصبونية أن تجد دالة رياضية بحيث تصل إلى النمط المستهدف بناء على النمط المدخل.أو بطريقة التعليم غير المراقب Unsupervised Learning وفيها يعرض مجموعة من الأنماط على الشبكة التي تقوم بضبط ذاتها لحل المشكلة المطروحة عليها بناء على مجموعة التدريب (Tylor B., 2006, P.2).

◄ **مكوّنات الشبكة العصبونية الصناعية Elements of ANN:**

تعتبر الخلية العصبونية الحجر الأساس للشبكة العصبونية الصناعية، فهي تتكون من مجموعة من عناصر المعالجة PEs تتوزّع فيما بينها ضمن مجموعة من الطبقات بحسب نوع الشبكة، تتصل مع بعضها البعض من خلال الأوزان (Kriesel D., 2005, p.33)، وكل عنصر معالجة يتألف من المدخلات Inputs وتمثل إما متغيرات الدراسة أي البيانات الخام أو مخرجات عناصر معالجة أخرى وتمثل على شكل شعاع X، والمخرجات Outputs قد تكون هي النتيجة النهائية للمهمة التي صمّمت الشبكة لأجلها أو مدخلات لعناصر معالجة أخرى وتمثل على شكل شعاع O، بالإضافة إلى دوال رياضة مسؤولة عن التجميع والاستثارة.

الشكل رقم (1)

بنية الخلية العصبونية رياضياً



المصدر: Mehrotra K,et al., Element of Artificial Neural Networks, P.58

بعد الانتهاء من إعداد الشبكة العصبونية تتم عملية معالجة المعلومات والتي تشمل (طلبة، وآخرون، 1998، ص390 – 393):

– مدخلات ومخرجات الشبكة Inputs and Outputs: وتمثل مدخلات الشبكة كامل متغيرات الدراسة، في حين أن مخرجاتها تمثل حل المشكلة التي صُممت الشبكة لأجلها.

– الأوزان Weights: وهو العنصر الأساسي في الشبكات العصبونية، ويعبر عن القوة النسبية للمتغيرات الداخلة في الشبكة أو عن قوة الربط بين طبقات الشبكة.

– دالة الجمع Summation Function: تقوم هذه الدالة بحساب وزن كامل المدخلات إلى عنصر المعالجة الواحد من خلال العلاقة التالية:

$$net = \sum_{i=1}^{n} x_i w_i$$

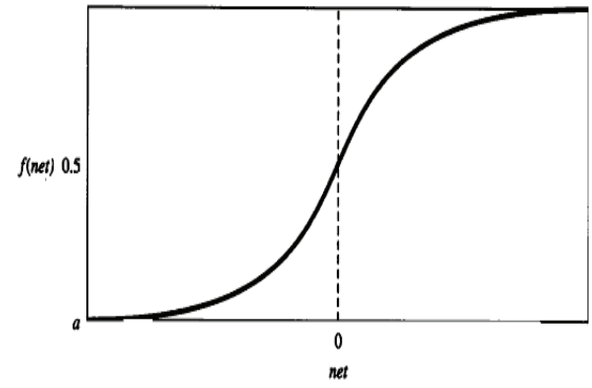
– دالة التفعيل Activation Function: لكل خلية عصبونية مستوى استثارة خاص بها، بعد أن تقوم دالة الجمع بعملها تستثار الخلية العصبونية وقد تعطي خرجاً معيناً، هذه العلاقة (بين دالة الجمع وخرج الخلية) قد تكون خطّية Linear أو لا خطّية Non Linear وتمثل عن طريق دالة التفعيل، هناك عدة أشكال لهذه الدالة يعتمد اختيار أحدها على طبيعة المشكلة التي تحلها الشبكة وعلى الباحث ذاته، إذ أثبتت معظم الدراسات أن دالة سيغمويد Sigmoid Function هي الأكثر كفاءة.

– دالة سيغمويد المنطقية Sigmoid Function: وهي الدالة الأكثر شهرة في الشبكات العصبونية تهدف إلى جعل قيمة المخرجات محصورة بين 0 و1 بسبب أن قيم المخرجات من ممكن أن تكون كبيرة جداً، وفيها يكون:

$$f(net) = \frac{1}{1 + e^{-net}}$$

$$\lim_{net \to -\infty} f(net) = 0 \qquad ، \qquad \lim_{net \to \infty} f(net) = 1$$

الشكل رقم (2)
التمثيل الرياضي للدالة المنطقية



المصدر: Mehrotra K.and others, Element of Artificial Neural Networks, P.14

◄ **الشبكات متعدّدة الطبقات** Multilayer Networks:

إن المشاكل المعقدة لا يمكن حلها بشبكات وحيدة الطبقة Perceptron، وظهرت بدلاً عنها شبكات التغذية الأمامية متعددة الطبقات Feedforward Multilayer بدوال غير خطية لعقدها، وبخوارزميات تعليم تختلف عنها.ظهرت هذه الشبكات في بداية الثمانينيات واعتمدت آلية تعليمها على الأخطاء للانتشار الخلفي Backpropagation أي من العقد الخلفية إلى العقد الأمامية.

◄ **معمارية الشبكات متعدّدة الطبقات** Architecture: تتوزع العقد في هذه الشبكات على عدد من الطبقات ترقم من 0 وحتى L بحيث أن الطبقة رقم 0 هي طبقة المدخلات والطبقة رقم L هي طبقة المخرجات، والرقم يدل على بعد عقد الطبقة عن عقد طبقة المدخلات، تطبق Backpropagation في حال كانت الشبكة تحوي على ثلاث طبقات على الأقل أي إذا كانت $L \geq 2$، وكل عقدة من العقد تمثل جزءاً من المشكلة المطروحة، فتمثل عقد طبقة المدخلات عدد أبعاد الأنماط المدخلة للشبكة أي متغيرات الدراسة، وتمثل عقد طبقة المخرجات عدد الأصناف بحسب المشكلة المطروحة، في حين أن عقد الطبقات المخفية تتحدّد بحسب طبيعة المشكلة ودرجة تعقيد الشبكة.يكون الاتصال بين العقد المتجاورة فقط أي لا يمكن لأي عقدة بأحد الطبقات أن تتصل بعقدة أخرى في طبقة غير مجاورة لها.

تطبّق كل عقدة موجودة في الطبقات المخفية وطبقة المخرجات الدالة المنطقية Sigmoid على مدخلاتها، وسبب اختيار هذه الدالة كونها دالة مستمرة Continuous ومطردة بالزيادة Monotonically Increasing، بالإضافة إلى مقارباتها عند أطراف اللانهايات، وحصر نتيجة تطبيقها ضمن المجال $[0,1]$.

الأهداف Objectives: تهدف الشبكة إلى جعل شعاع المخرجات بعد التدريب op قريباً جداً من شعاع المخرجات الفعلي dp، أي تكون قيمة دالة التقييم أقل ما يمكن:

$$Error = \sum_{p=1}^{P} Err(o_p, d_p) \rightarrow Min$$

ويمكن اعتماد مجموع مربعات الأخطاء في تقييم نتيجة الشبكة:

$$SSE = \sum_{\text{عينة}=1}^{P} \sum_{i=1}^{K} (\ell_{p,j})^2$$

$$\ell_{p,j} = |o_{p,j} - d_{p,j}|$$

خوارزمية الانتشار الخلفي Backpropagation Algorithm (Mehrotra, et al., 1997, P.70 - 79): تسعى هذه الخوارزمية لتخفيض قيمة الخطأ بين مخرجات الشبكة والمخرجات المطلوبة من خلال تعديل الأوزان، وتعتبر من خوارزميات التعليم المراقب Supervised، تقوم شبكة Feedforward بتقديم الأنماط المدخلة إلى عقد طبقة المدخلات التي بدورها تقوم بتمريرها لعقد الطبقة المخفية الأولى، إن العقد في كل طبقة مخفية تحسب مجموع الأوزان لمدخلاتها ومن ثم تطبق عليها دالة التفعيل وتمرّرها للطبقة التي تليها حتى الوصول إلى طبقة المخرجات.

يتم ترميز كامل الأوزان في الشبكة بمصفوفة W ويكون هدف الخوارزمية هو إيجاد قيمة الأوزان المناسبة بحيث تكون قيمة دالة الخطأ أقل ما يمكن.يقصد بالرمز الوزن المسند للعلاقة بين العقدة j في الطبقة i مع العقدة k في الطبقة i+1، ويقصد بالرمز مخرجات العقدة j في الطبقة i للنمط p، ويحسب الخطأ لكل نمط p مدخل

بالعلاقة:

$$E_p = \sum_k (\ell_{p,k})^2$$

وبإزالة أثر النمط p من الدالة السابقة يصبح لها الشكل التالي:

$$E_p = \sum_{k=1}^{K} (\ell_k)^2$$

ويكون الهدف هو تخفيض قيمة الخطأ، ووفقاً لطريقة الانحدار Gradient Descent Method فإن الأوزان تتغير باتجاه يتوافق مع اتجاه مشتق الخطأ بالنسبة للوزن $\partial E/\partial w$. وبتطبيق قاعدة دلتا المعمّمة Generalized Delta Rule ممكن تسهيل حساب المشتق السابق، بالاعتماد على قاعدة التسلسل في الاشتقاق The Chain Rule of Derivatives وافتراض أن الشبكة تحوي على طبقة مخفية واحدة فقط، يمكن تجزئة تغيرات الوزن بدراسة المشتق من الطبقة المخفية إلى طبقة المخرجات $\partial E/\partial w_{k,j}^{(2,1)}$ ودراسة المشتق من طبقة المدخلات إلى الطبقة المخفية $\partial E/\partial w_{j,i}^{(1,0)}$.

وبعد إجراء العمليات الرياضية نحصل على الدالة التالية لتعديل الأوزان بين عقد الطبقة المخفية وعقد طبقة المخرجات:

$$\Delta w_{k,j}^{(2,1)} = \eta\,(d_{p,k} - o_{p,k})\,\delta'(net_{p,k}^{(2)})\,x_{p,j}^{(1)}$$

والدالة التالية لتعديل الأوزان بين عقد طبقة المدخلات وعقد الطبقة المخفية:

$$\Delta w_{j,i}^{(1,0)} = \eta \sum_k \left((d_{p,k} - o_{p,k})\,\delta'(net_{p,k}^{(2)})\,w_{p,j}^{(2,1)} \right) \delta'(net_{p,j}^{(1)})\,x_{p,i}$$

◄ **مفهوم تحليل التمايز Discriminant Analysis:**

يعتبر من أساليب تحليل المتغيرات المتعددة التي تهتم بفصل عدد من المشاهدات إلى مجموعتين أو أكثر وتوزيع المشاهدات الجديدة عليها، فهو يستخدم لاستكشاف أسباب الاختلافات المشاهدة عندما لا نستطيع فهم العلاقات السببية بدرجة كافية (ريتشارد، وشرن، 1998، ص 740 – 735)، ويمكن تعريفه على أنه «أسلوب احصائي لبناء نموذج رياضي يكون المتغير التابع فيه ممثلاً برموز (رمزين على الأقل) كل رمز يمثل رقماً لمجموعة من المجموعات الداخلة في التحليل، ويكون بالنموذج بعض المتغيرات المستقلة» (عاشور، وآخرون، 2004، ص140)، ويستخدم النموذج للتنبؤ برقم المجموعة التي تنتمي إليها المشاهدة الجديدة من خلال المتغيرات المستقلة الموجودة في النموذج.

يوجد نوعان أساسيان لتحليل التمايز هما: تحليل التمايز الخطي LDA (Linear Discriminant Analysis)وهو تحليل التمايز البسيط الذي يصنف المشاهدات إلى مجموعتين فقط؛ أي أن المتغير التابع هو متغير ثنائي التصنيف، ويرتبط هذا النوع بشكل وثيق مع تحليل التباين ANOVA وتحليل الانحدار البسيط.وتحليل التمايز المتعدد MDA (Multiple Discriminant Analysis)وفيه يكون المتغير التابع له أكثر من قيمتين ويرتبط بتحليل التباين المتعدد MANOVA وتحليل الانحدار المتعدد (Geoffry, Lachlan, 1992).

في حالة تحليل التمايز الخطي يكون لدينا دالة تصنيفية واحدة فقط، وفي حالة تحليل التمايز المتعدد يكون لدينا أكثر من دالة تصنيفية يصل عدد كحد أقصى ل (عدد المتغيرات المستقلة – 1)، ويكون لدالة التمايز الشكل التالي:

$$D_t = \lambda_{t0} + \lambda_{t1}x_1 + \lambda_{t2}x_2 + \cdots + \lambda_{tk}x_k$$

حيث:

- Dt: نتيجة التمايز بين مجموعتين محددتين.
- t: عدد دوال التمايز.
- K: عدد المتغيرات المستقلة.
- $\lambda_{t1}, \lambda_{t2}, \lambda_{tk}$: معاملات المتغيرات المستقلة على الترتيب.

شروط تطبيق تحليل التمايز البسيط: يرى (العدوي، 2017) أن أهم شروط يجب أن تتوافر لتطبيق تحليل التمايز البسيط هي:

- وجود قيمتين فقط للمتغير التابع.
- يجب أن يكون حجم العينة كبيراً، لكل متغير 20 مشاهدة على الأقل.
- عدم وجود مصاحبة خطية بين المتغيرات المستقلة، أي يجب أن تكون مستقلة فيما بينها.
- عدم وجود قيم متطرفة للمتغيرات.
- المجتمعات الإحصائية تمثل مجتمعات إحصائية ذات توزيع طبيعي.

الدراسة التطبيقية:

يقوم النموذج المقترح على دراسة العوامل المؤثرة بشكل أساسي في اتخاذ قرار الكفالة المالية، ومن ثم تقييم كل حالة لدعم قرار تصنيفها إلى فئة المقبولين للكفالة أم لا، مما يساهم في مساعدة متخذ القرار لتصنيف طالبي الكفالة بالشكل الأمثل.

طبقت الدراسة على قاعدة بيانات عدد من الجمعيات الخيرية العاملة في مدينة حلب – سورية والتي تحوي على قسم مختص بالكفالات المالية خلال الفترة (من عام 2016 وحتى عام 2019)، تحوي قاعدة البيانات على 12 متغيرا منها 4 متغيرات مستمرة و4 متغيرات ثنائية و5 متغيرات فئوية، ووجود 7463 حالة بدون وجود قيم مفقودة وهي أعلى من حجم العينة 6344 اللازم لإجراء الدراسة.وفيما يلي توضيح بالمتغيرات مع الترميز الخاص بكل متغير:

الجدول رقم (1)

متغيرات الدراسة

المتغيرات المستمرة	
العمر	X1

المتغيرات المستمرة		
عدد أفراد الأسرة الكلي		X3
عدد الأفراد الممكن كفالتهم		X4
عدد الأفراد المكفولين فعلاً		X5
المتغيرات الثنائية (0 تعني لا، 1 تعني نعم)		
الجنس (ذكر – أنثى)		X2
هل الأم على قيد الحياة		X7
هل يوجد حالة إعاقة		X8
هل تتم الكفالة		Y
المتغيرات الفئوية		
تقييم العائلة	جيد – متوسط – سيء – غير محدد	X6
حالة العائلة	مقيم – نازح – متضرر	X9
نوع السكن	ملك – آجار – مستضاف – غير ذلك	X10
حالة السكن	مفروش – مكسي – هيكلي – غير محدد	X11
نوع طالب الكفالة	يتيم – طالب جامعي – محتاج	X12

المصدر: من إعداد الباحثين بالاعتماد على عينة الدراسة

تطبيق الشبكة العصبونية على البيانات:

إن بنية الشبكة العصبونية التي سيعتمدها الباحث لتناسب مسألة التصنيف هي شبكة متعددة الطبقات Multilayer Layer Per-ceptron (MLP)، سيتم تحديد عناصر هذه الشبكة وفق ما يلي:

- طبقة المدخلات Input Layer: تتمثل إجمال المتغيرات الداخلة في الدراسة وعددها 13 متغيرا، أي يوجد 13 عقدة في طبقة المدخلات.
- طبقة المخرجات Output Layer: تمثل متغير القرار، وهو المتغير الذي قسم الأفراد إلى مجموعتين أساسيتين مقبول وغير مقبول، وبالتالي فإن عدد عقد طبقة المخرجات هو عقدتان.
- الطبقات المخفية Hidden Layer: حدد الباحث طبقة مخفية واحدة فقط للشبكة العصبونية بعد الاعتماد على برنامج SPSS في تحديد عدد الطبقات بشكل أوتوماتيكي، وترك عدد العقد للبرنامج ضمن الطبقة، إذ يعتمد النظام على مبدأ التجريب لحين الوصول إلى عدد عقد يعطي حلا جيدا.

طريقة قياس المتغيرات: حدد الباحث الطريقة المعيارية Standared التي تعتمد على الوسط الحسابي Mean والانحراف المعياري Standard Deviation في حساب القيم.

عينتا التدريب والاختبار Training and Testing Set: حدد الباحث نسبة 80% من المفردات لتدخل في عينة التدريب دون أية شروط، وتبقى نسبة 20% لتشكل مفردات عينة الاختبار.

دوال التفعيل Activation Function: اعتمد الباحث على الدالة Sigmoid في عقد كل من طبقتي المخرجات والخفية.

نمط التدريب Type of Training: حدد الباحث طريقة تحديث الأوزان بعد إدخال كامل الأنماط، أي استخدم طريقة P - E بسبب الحاجة إلى السرعة في المعالجة.

خوارزمية تحقيق المثلوية Optimization Algorithm: حدد الباحث طريقة الانحدار Gradiant Descent المترافقة مع خوارزمية الانتشار الخلفي Backpropagation، وقد حدد الباحث وسطاء الخوارزمية وفق التالي:

- قيمة معدل التعليم Learning Rate: η = 0.4.
- قيمة وسيط عزم التحريك Momentum: α = 0.9.
- القيم الأولية للأوزان Initialization of Weights: محصورة ضمن المجال [– 0.5،0.5+].

ضبط شروط توقف الخوارزمية: تقف الخوارزمية عند الوصول إلى تكرار لم تتغير فيه قيمة مجموع مربعات الأخطاء SSE، وفي حال لم نصل لهذا الشرط تتوقف الخوارزمية بعد 15 دقيقة.

نتائج تنفيذ الشبكة العصبونية في SPSS:

بعد ضبط عناصر الشبكة حصل الباحث على النتائج الآتية:

تقسيم البيانات على العينات: قسمت البيانات وفقاً للنسبة التي حددها الباحث كما هي في الجدول الآتي:

الجدول رقم (2)

موجز البيانات Case Processing Summary

	العدد N	النسبة %
مفردات عينة التدريب	5995	80.3%
مفردات عينة الاختبار	1468	19.7%

المصدر: من إعداد الباحثان بالاعتماد على مخرجات SPSS

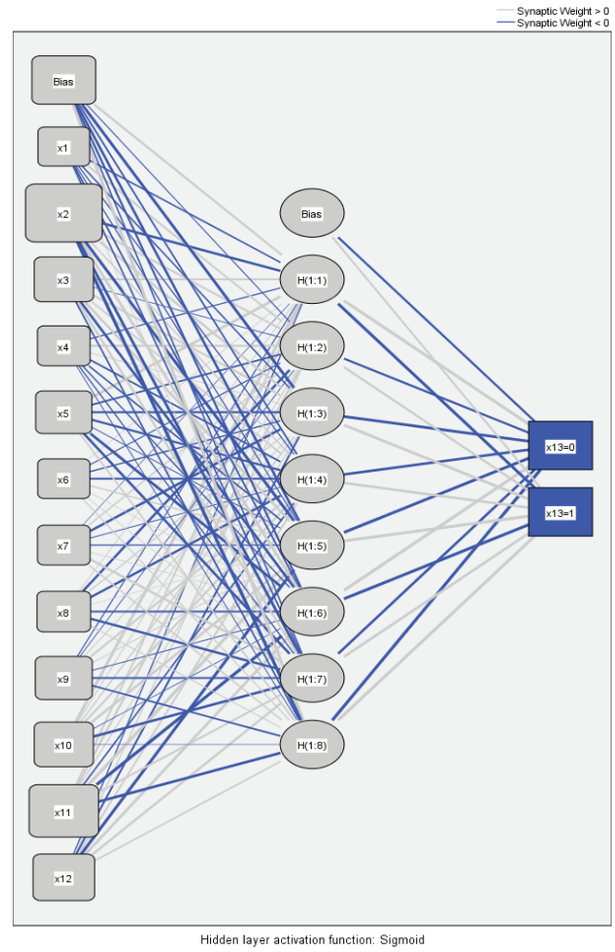
معلومات الشبكة Network Information: تحوي الشبكة على ثلاث طبقات، تحوي الطبقة الأولى على 12 عقدة هي المتغيرات اللازمة لاتخاذ القرار بالإضافة إلى متغير بييز Bais وقد اعتمد في حسابها على الطريقة المعيارية Standardized، بالإضافة إلى 9 ققد في الطبقة المخفية بما فيها متغير Bais ودالة تفعيلها هي Sig-moid، وعقدتين في طبقة المخرجات بدالة تفعيل Sigmoid.ودالة الخطأ هي مجموع مربعات الأخطاء Sum of Squares، والشكل رقم (3) يوضح بنية الشبكة العصبونية الناتجة من التحليل.

أداء الشبكة Network Performance:

تقاس من خلال مجموعتي التدريب والاختبار معاً، ومن خلالها نحكم على جودة الشبكة، وتعطي النتائج التالية:

الشكل رقم (3)

بنية الشبكة العصبونية الناتجة عن التحليل



المصدر: من إعداد الباحثان بالاعتماد على مخرجات SPSS

ملخص النموذج Model Summary: إن مجموع مربعات الأخطاء بالنسبة لعينة التدريب هو 371.571 بنسبة خطأ للتقدير 7.9%، والمدة التي استغرقته الشبكة للتدريب هي 72 جزء من الثانية، وبالنسبة لمجموع مربعات الأخطاء بالنسبة لعينة الاختبار فقد بلغت 83.732 بنسبة خطأ للتقدير 7.2% كما هو موضح في الجدول التالي:

الجدول رقم (3)

ملخص النموذج Model Summary

عينة التدريب		عينة الاختبار	
SSE	371.571	SSE	83.732
نسبة التصنيف الخاطئ	7.9%	نسبة التصنيف الخاطئ	7.2%
زمن التدريب: 0: 00: 00: 72			

المصدر: من إعداد الباحثان بالاعتماد على مخرجات SPSS

تصنيف مفردات العينة Classification: إن مفردات التدريب قد صنفت بشكل صحيح بنسبة 92.1%، أما بالنسبة لمفردات الاختبار فتم تصنيفها بشكل صحيح وبنسبة 92.8%، كما هو واضح في الجدول التالي:

الجدول رقم (4)

تصنيف مفردات العينة Classification

Classification

Sample	Observed	Predicted 0	Predicted 1	Percent Correct
Training	0	826	333	71.3%
	1	142	4678	97.1%
	Overall Percent	16.2%	83.8%	92.1%
Testing	0	194	74	72.4%
	1	33	1183	97.3%
	Overall Percent	15.3%	84.7%	92.8%

Dependent Variable: مكفول هو هل

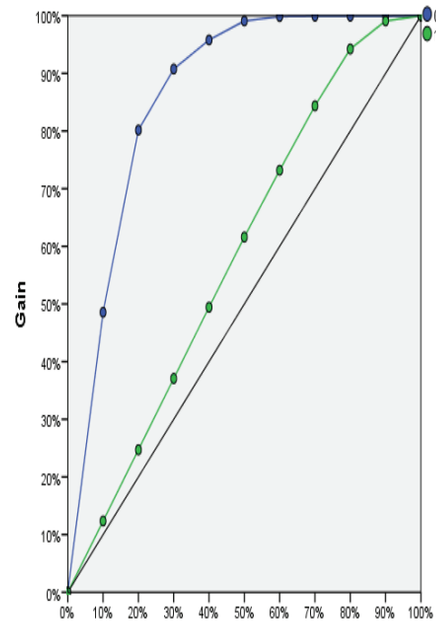
المصدر: من إعداد الباحثان بالاعتماد على مخرجات SPSS

مخطط العائد التجميعي Cumulative Gains Chart: يعرض هذا المخطط نسبة العدد الإجمالي للحالات التي كسبتها الفئة الواحدة، من خلال استهداف نسبة العدد الإجمالي للحالات، ونلاحظ من المخطط أن النقطة الأولى لمنحنى لفئة 1 (ذو اللون الأزرق) تقع في الإحداثيات (10%، 50%)، وهذا يعني التوقع بأن أعلى 10% من الحالات تحوي ما يقارب على 50% من الحالات التي تأخذ فعلياً القيمة 0، وبالمثل فإن 20% من الحالات تحوي ما يقارب 80% وهكذا...حتى الوصول إلى 100% من الحالات والعائد. يمثل الخط المائل منحنى خط الأساس Baseline Curve الذي يعني بأن 10% من الحالات تقابل 10% من العائد المتوقع التي تأخذ القيمة نعم، وبما أن كامل النقط أعلى من منحنى خط الأساس فالشبكة جيدة في التنبؤ للحالات التي تكون نتيجتها 0، وللحالات التي تكون نتيجتها 1، ويظهر هذا المخطط وفق الشكل رقم (4).

أهمية المتغيرات المستقلة Independent variable importance: وهو مقياس لمقدار توقع نموذج الشبكة لتغيرات القيم وفقاً للمتغيرات المستقلة، وتكمن أهمية القيم المطبّعة Normalized بحساب النسبة المئوية لكل متغير من خلال قسمة قيمته على أعلى قيمة موجودة كما هو موضح في الجدول رقم (5) :

الشكل رقم (4)

مخطط العائد التجميعي Cumulative Gains Chart



المصدر: من إعداد الباحثان بالاعتماد على مخرجات SPSS

الجدول رقم (5)

أهمية المتغيرات المستقلة Independent variable importance

Independent Variable Importance

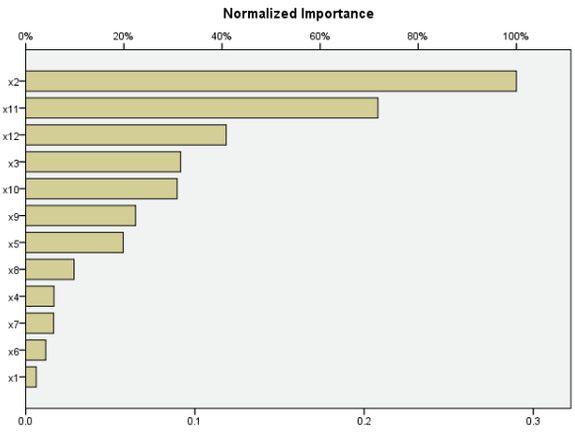
	Importance	Normalized Importance
الجنس	.006	2.1%
العمر	.290	100.0%
المكفول نوع	.092	31.6%
الأفراد عدد	.017	5.8%
القديم	.058	19.9%
الأم وجود	.012	4.1%
المنزل نوع	.017	5.7%
المنزل حاله	.029	9.8%
الاعاقات عدد	.065	22.4%
العائله نوع	.089	30.8%
المكفولين عدد	.208	71.8%
الأيتام عدد	.118	40.9%

المصدر: من إعداد الباحثان بالاعتماد على مخرجات SPSS

وكما هو واضح فإن العمر ذو أهمية كبيرة في النموذج يليه عدد المكفولين وهكذا...، لحين الوصول إلى أقل المتغيرات أهمية وهو متغير الجنس.والشكل التالي يوضح أهمية هذه المتغيرات:

الشكل رقم (5)

أهمية المتغيرات المستقلة Independent variable importance

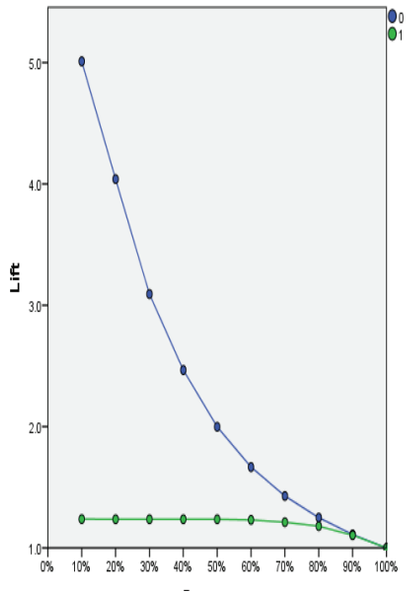


المصدر: من إعداد الباحثان بالاعتماد على مخرجات SPSS

مخطط LIFT: هذا المخطط ناتج عن المخطط السابق، وفيه تكون القيم على المحور العمودي ناتجة من خلال قسمة قيمة الكسب التجميعي على الخط الأساس Base Line، لذا نجد أن القيمة عند 10% هي ناتج قسمة 50% على 10% وهي 5، فهو شكل آخر للمنحنى السابق.

الشكل رقم (6)

مخطط LIFT



المصدر: من إعداد الباحثان بالاعتماد على مخرجات SPSS

قوة العلاقة بين عقد الطبقات Parameter Estimates: من خلال هذا الجدول يتبين قوة العلاقة بين كل عقدة من عقد طبقة المدخلات مع كل عقدة من عقد الطبقة المخفية، وقوة العلاقة بين كل عقدة من عقد الطبقة المخفية مع كل عقدة من عقد طبقة المخرجات، وواضح أن الاتصال بين عقدة عدد المكفولين x11 من طبقة المدخلات وباقي العقد في الطبقة المخفية أقوى من غيرها من الاتصالات بين العقد وهذا يؤكد ما أظهرته نتيجة أهمية العوامل المستقلة، والجدول التالي يوضح كامل هذه القيم:

الجدول رقم (6)

قوة العلاقة بين عقد الطبقات Parameter Estimates

Parameter Estimates

Predictor		Predicted									
		Hidden Layer 1								Output Layer	
		H(1:1)	H(1:2)	H(1:3)	H(1:4)	H(1:5)	H(1:6)	H(1:7)	H(1:8)	[x13=0]	[x13=1]
Input Layer	(Bias)	.592	-.108	-1.898	-.301	-.912	-.462	-.211	2.161		
	x1	-.261	.437	-.228	-.141	-.132	-.303	-.239	.011		
	x2	-.985	.458	.667	.209	-2.407	1.586	-1.793	-5.517		
	x3	.485	-.048	.536	.308	.744	-1.069	.544	-.292		
	x4	-.143	.524	.057	-.309	-.270	-.180	-.157	-.058		
	x5	1.859	-.315	-.572	-.538	-.058	-1.606	-.982	.298		
	x6	.011	-.122	-.206	-.444	.202	.010	.028	.140		
	x7	-.001	.197	-.919	-.069	-.163	.020	.693	.218		
	x8	.043	-.704	.186	-.686	.008	-.382	-.705	.108		
	x9	-.111	-.051	.095	.530	.095	-.081	-.413	-.327		
	x10	2.600	.072	.206	.519	1.045	-.033	-1.193	-.003		
	x11	.392	.676	.121	-.246	4.585	-5.281	.460	-1.827		
	x12	-.144	.152	-.413	.333	-3.039	1.502	.677	.298		
Hidden Layer 1	(Bias)									-.603	.449
	H(1:1)									3.192	-3.105
	H(1:2)									-.621	.674
	H(1:3)									-2.216	2.265
	H(1:4)									-1.227	1.166
	H(1:5)									-3.329	3.434
	H(1:6)									3.678	-3.568
	H(1:7)									-1.155	1.207
	H(1:8)									-3.629	3.606

المصدر: من إعداد الباحثان بالاعتماد على مخرجات SPSS

تطبيق تحليل التمايز:

تم تطبيق تحليل التمايز على كامل أفراد العينة ولكل المتغيرات، وبما أن حجم العينة كبير 7463 يمكن إهمال شرط تجانس التباين، وتم تطبيق اختبار Wilks Lambda للتحقق من فرضية عدم تساوي متوسطات المتغيرات المستقلة عند مستويات المتغير التابع حيث يجب أن يكون قيمة Sig لكل متغير مستقل أقل من 05.، نلاحظ أن كلا من الجنس ونوع المكفول وعدد الأفراد ووجود الأم غير معنويات من الناحية الإحصائية، حيث كانت قيمة Sig لكل منها أكبر من 05.، أما بقية المتغيرات فهي معنوية وذات دلالة إحصائية إذ كانت قيمة Sig لكل منها أقل من 05..

وعند القيام بإجراء تشخيص المصاحبة الخطية على أبعاد المتغير المستقل نجد أنّ قيمة السماحية Tolerance (تعبر عن مقدار تباين المتغير المستقل المحدد الذي لا تفسره المتغيرات المستقلة الأخرى في النموذج ويحسب من خلال $1 - R^2$) مرتفعة لكل المتغيرات المستقلة (لا تقل عن 10.) وبالتالي لم يحدث إخلال بفرضية المصاحبة الخطية، ويدعم هذه النتيجة أيضاً قيمة معامل تضخم التباين VIF والذي يبين أن جميع القيم أقل من 10.والجدول التالي يبين قيم السماحية ومعامل تضخم التباين وإحصائية الاختبار وقيمة Sig لكل متغير مستقل:

الجدول رقم (7)

المصاحبة الخطية و إحصاءات المتغيرات المستقلة

المتغير	VIF	Tolerance	Wilks' Lambda	Sig.
الجنس	1.005	.995	1.000	.525
العمر	1.507	.664	.933	.000
نوع المكفول	1.659	.603	1.000	.174
عدد الأفراد	2.122	.471	1.000	.329
التقييم	1.299	.770	.961	.000
وجود الأم	1.066	.938	1.000	.272
نوع المنزل	1.108	.903	.998	.000
حالة المنزل	1.111	.900	.999	.003
عدد الإعاقات	1.032	.969	.999	.016
نوع العائلة	1.312	.762	.991	.000
عدد المكفولين	1.476	.678	.868	.000
عدد الأيتام	2.316	.432	.999	.014

المصدر: من إعداد الباحثان بالاعتماد على مخرجات SPSS

تم تطبيق اختبار Wilks Lambda للتحقق من معنوية النموذج بعد استبعاد المتغيرات المستقلة غير المعنوية، إذ تنص فرضية العدم على أن النموذج غير معنوي، ونلاحظ بأن قيمة Sig مساوية للصفر وهي أقل من 05. لذلك نرفض فرضية العدم ونقبل الفرضية البديلة والنموذج معنوي من الناحية الإحصائية، ونلاحظ أن قيمة معامل التحديد للنموذج 55% وهي تمثل النسبة التي استطاعت المتغيرات المستقلة مجتمعة أن تفسره من تغيرات المتغير التابع، وقيمة معامل الارتباط المتعدد للنموذج هي 74.1% وهي قيمة جيدة، كما هو مبين في الجدول التالي:

الجدول رقم (8)

ملخص نموذج تحليل التمايز

اختبار Wilks' Lambda			الجذر الكامن		
القيمة	df	Sig	Eigenvalue	R^2	R
.697	8	.000	.434	.550	.741

المصدر: من إعداد الباحثان بالاعتماد على مخرجات SPSS

وتكون دالة التمايز لها الشكل التالي:

D=-2.199+0.070 X1+0.3 X2+0.089 X3+0.181 X4-0.092 X5+0.407 X6-0.761 X7+0.47 X8

حيث أن:

X1	العمر	X5,	عدد الاعاقات
X2	التقييم	X6,	نوع العائلة
X3	نوع المنزل	X7,	عدد المكفولين
X4	حالة المنزل	X8,	عدد الأيتام

والجدول التالي يبين أهمية المتغيرات المستقلة:

الجدول رقم (9)

أهمية المتغيرات المستقلة

Structure Matrix

	Function
	1
المكفولين عدد	-.586
العمر	.403
التقييم	.302
العائلة نوع	.142
المنزل نوع	.070
المنزل حالة	.052
الأيتام عدد	-.043
الاعاقات عدد	.042
المكفول نوع	.024
الأم وجود	.019
الأفراد عدد	.017
الجنس	-.011

المصدر: من إعداد الباحثان بالاعتماد على مخرجات SPSS

نلاحظ من الجدول السابق أن عدد المكفولين له التأثير الأكبر في النموذج، يليه العمر ومن ثم التقييم...وهكذا، ونلاحظ أن الجنس وعدد الأفراد ووجود أم ونوع المكفول ليس لها أي تأثير على النموذج.

وكانت دقة التصنيف للدالة التمييزية هي:

$$C = \frac{1155 + 4892}{1155 + 272 + 1144 + 4892} = \frac{6047}{7463}$$

$$C = 0.81 = 81\%$$

أي استطاعت الدالة تمييز 81% بشكل صحيح من إجمالي الحالات المدروسة، كما هي مبينة في الجدول التالي:

الجدول رقم (10)

التصنيف باستخدام تحليل التمايز Classification

هل هو مكفول أم لا		نتيجة الدالة	
		0	1
القيم الأصلية	0	1155	272
	1	1144	4892

المصدر: من إعداد الباحثان بالاعتماد على مخرجات SPSS

مقارنة نتائج الشبكة العصبونية وتحليل التمايز:

إن نسبة التصنيف الصحيحة في تحليل التمايز أقل مما نتج عن الشبكة العصبونية، حيث يبين الجدول التالي نسبة التصنيف بحسب كل طريقة:

الجدول رقم (11)

مقارنة نتيجة التصنيف

الأسلوب	نسبة التصنيف الصحيحة	نسبة التصنيف الخاطئة
تحليل التمايز	81%	19%
الشبكة العصبونية/ التدريب	92.1%	7.9%
الشبكة العصبونية/ الاختبار	92.8%	7.2%

المصدر: من إعداد الباحثان بالاعتماد على مخرجات SPSS

ونلاحظ من النتائج السابقة تفوق الشبكة العصبونية في كلا المرحلتين (التدريب والاختبار) على تحليل التمايز، كما نلاحظ بأن النسبة الصحيحة للتصنيف مقبولة مما يؤدي إلى إمكانية اعتماد هذه الشبكة للمساعدة في اتخاذ القرار اللازم بالكفالة المالية في الجمعيات الخيرية.

النتائج:

1. لقد أعطى استخدام الشبكات العصبونية في التصنيف نتيجة أفضل من تنفيذ تحليل التمايز على ذات البيانات لغرض اتخاذ قرار تصنيف الكفالة المالية، إذ نجد أن دقة التصنيف في الشبكة العصبونية 92.8% في حين أن نسبة التصنيف الصحيح في أسلوب تحليل التمايز 81%.
2. إمكانية تدريب الشبكة العصبونية على الأنماط الجديدة بشكل ديناميكي متفوقة بذلك على تحليل التمايز.
3. اعتمدت الشبكة العصبونية الصناعية على كامل

المتغيرات للتنبؤ بقرار التصنيف، في حين أن تحليل التمايز استثنى أربعة متغيرات لعدم معنويتهم في النموذج.

4. إن المتغيرات الأكثر تأثيراً في اتخاذ قرار تصنيف الكفالة المالية هي العمر وعدد المكفولين في كلا الأسلوبين.

5. إن المتغيرات الأضعف لاتخاذ قرار تصنيف الكفالة المالية باستخدام الشبكات العصبونية هي الجنس ووجود الأم ونوع المنزل.

رؤية تطبيقية للنتائج:

ينصح الجمعيات الخيرية باستخدام الشبكات العصبونية الصنعية لغرض اتخاذ قرار تصنيف الكفالة المالية، لدقتها وتكيفها مع الحالات الجديدة واعتمادها على كامل البيانات. والاهتمام بالمتغيرات الأكثر تأثيراً في اتخاذ قرار تصنيف الكفالة المالية وهي العمر وعدد المكفولين.

التوصيات لأبحاث مستقبلية:

1. دراسة تغيير بنى الشبكات العصبونية من حيث ثوابت الشبكة والخوارزميات الخاصة بالتدريب وتوابع التفعيل في مختلف الطبقات على نتيجة التصنيف.

2. دراسة النماذج الإحصائية الأخرى مثل نموذج الانحدار اللوجستي مع الشبكات العصبونية لاتخاذ قرار التصنيف في حالة البيانات المتعددة.

3. تطوير الشبكة العصبونية بإدخال التحليل العاملي على طبقة المدخلات وبالتالي تخفيض أبعاد الدراسة إلى عدد أقل من العوامل، ودراسة أثر ذلك على اتخاذ القرار.

4. مقارنة الشبكة العصبونية متعددة الطبقات مع الخرائط ذاتية التنظيم SOM في التصنيف.

5. الانتقال من الأسلوب التقليدي في عملية اتخاذ القرار إلى اعتماد نماذج حديثة كالذكاء الصنعي في عملية اتخاذ القرار.

المصادر والمراجع باللغة العربية:

- العدوي، عدي. (2017)، دراسة نماذج الانحدار للمتغيرات النوعية، رسالة ماجستير غير منشورة، كلية العلوم، جامعة حلب، سوريا، ص24.
- جونسون، ريتشارد. وشرن، دين. (1998)، التّحليل الاحصائي متعدد المتغيّرات من الوجهة التطبيقية. تعريب عبد المرضي عزام، دار المريخ، الرياض، ص735 – 740.
- سليمان، علي أبشر فضل المولى. (2015)، المقارنة بين التحليل التمييزي والنموذج اللوجستي الثنائي ونماذج الشبكات العصبية في تصنيف المشاهدات: بالتطبيق على دراسة العوامل المؤثرة على كفاية دخل الأسرة، أطروحة دكتوراه غير منشورة، كلية العلوم، جامعة السودان للعلوم والتكنولوجيا، السودان.
- طلبة، محمد، وآخرون. (1998)، الحاسب والذكاء الاصطناعي. موسوعة التميز لعلوم الإدارة والحاسب، القاهرة، مصر، ص390 – 393.
- عاشور، سمير كامل، وآخرون. (2004) العرض والتحليل الاحصائي باستخدام Spss Win. القاهرة، مصر، ص140، (د. د).

ترجمة المصادر المراجع العربية:

- *Adawi, O. (2017). Study of Regression Models of Qualitative Variables, Unpublished Master Thesis, Faculty of Science, University of Aleppo, Syria, p. 24.*
- *Johnson, R. and Schorn, D. (1998). Applied Multivariate Statistical Analysis. The translation of Abd al - Mardi Azzam, Dar al - Merikh, Riyadh, pp. 735 - 740.*
- *Solaiman, M. (2015). Comparison between discriminatory analysis, binary logistic model, and neural network models in rating views: by applying to a study of factors affecting family income adequacy, unpublished Phd thesis, Faculty of Science, Sudan University of Science and Technology, Sudan.*
- *Talaba, M. et al. (1998). Computer and Artificial Intelligence. Encyclopedia of Excellence in Management and Computer Sciences, Cairo, Egypt, pp. 390 - 393.*
- *Ashour, Samir Kamel, et al. (2004). Demonstration and statistical analysis using Spss Win. Cairo, Egypt, p. 140.*

المراجع باللغة الإنكليزية:

- *Barzanti Luca, Giove Silvio and Pezzi Alessandro. (2017) , A Decision Support System for Non Profit Organizations, Springer International Publishing.*
- *Geoffry. J, MC. Lachlan (1992) , Discriminant Analysis and Statistical Pattern Recognition; John Wiley & Sons, Inc.*
- *Khemakhem Sihem, Boujelbène Younés, "Credit risk prediction: A comparative study between discriminant analysis and the neural network approach", Accounting and Management Information Systems, Vol. 14, No. 1, University of Sfax, Tunisia, 2015.*
- *Kriesel D. (2005) , "A Brief Introduction to Neural Networks", available at [http: // www. dkriesel. com/ en/ science/ neural_ networks], Bonn, Germany, P33.*
- *Mehrotra K. , Mohan C. , Ranka S. (1997) , Element of Artificial Neural Networks, MIT Press, USA, P70 - 79.*
- *Leela Rani K, G Loshma, "Classification and Prediction in CRM Using Back Propagation Multilayer Feedforward Neural Network Approach", International Journal of Commerce, Business and Management, Vol. 1, No. 1, 2012, India.*
- *Tylor B. (2006) , Methods and Procedures for the Verification and Validation of Artificial Neural Networks, Springer, USA, P2.*