بناء شجرة القرار باستخدام خوارزمية C4.5 لدعم قرارات الجمعيات الخيرية (دراسة تطبيقية على جمعية أهل الخير بحلب)

Build a Decision Tree Using the C4.5 Algorithm to Support Charities Decisions Applied Study on the Charity Ahl al - khair in Aleppo

هبة محمد أمير زنرني
طالبة دكتوراه/ جامعة حلب/ سوريا

غسان ساكت
أستاذ دكتور/ جامعة حلب/ سوريا

Hiba Mohammed Amir Zannerni
PhD. Student/ University Aleppo/ Syria
hibazannerni@gmail.com

Ghassan Saket
Professor/ University Aleppo/ Syria
Ghsaket@hotmail.com

تاريخ الاستلام: 27/ 11/ 2019م، **تاريخ القبول:** 10/ 5/ 2020م.
E-ISSN: 2410-3349
P-ISSN: 2313-7592

Received: 27/ 11/ 2019, **Accepted:** 10/ 5/ 2020.
DOI: 10.33977/1760-005-013-011
https: //journals.qou.edu/index.php/eqtsadia

الملخص:

إنَّ عملية التقييم وتحديد المعايير المناسبة لاتخاذ القرار في الجمعيات الخيرية ليست من العمليات البسيطة، فمن خلالها يتم تحديد الاحتياج وتقييم وتصنيف العوائل وإعطاء القرار المناسب، وقد ظهرت ضمن فترة الأزمة ما بين عام 2012 إلى عام 2019 عدة مشاكل تحتاج لاتخاذ قرار سليم ومناسب ضمن الجمعيّات، وتحتاج مثل هذه القرارات إلى الدقة والوضوح لعدم حدوث ظلم أو تحيز أو خطأ أثناء التقييم.

في البداية قامت الجمعيّات الخيرية بتحديد المعايير اللازمة لاتخاذ القرار ومن ثم إيجاد آلية بسيطة لتسهيل عملية اتخاذ القرار، ولكن ما تزال نسبة الخطأ موجودة أثناء عملية تقييم العوائل، حيث اعتمد البحث على المنهج الوصفي وهدف إلى تحديد المعايير ودراستها وتحليلها لتصنيف العوائل بحسب الأولية، و درجة الاحتياج وذلك باستخدام خوارزمية C4.5، وتم استخدام الأداة WEKA في بناء النموذج لتقييم العوائل، فقد تم بناء شجرة القرار بالاعتماد على مجموعة من بيانات حجمها كبير، وتم تقييم نموذج شجرة القرار من خلال تفسير دقيق للقيم الناتجة.

وقد توصل البحث إلى إمكانيّة بناء نموذج شجرة قرار تصنيف العوائل ذات دقة عالية إذ بلغت دقة النموذج باستخدام خوارزمية C4.5 (0.9691)، وتم تقييم النموذج من خلال الأداة WEKA، وقد أوصت الباحثة بعد بناء النموذج بأهمية تطوير علمية اتخاذ القرار باستخدام تقنيات إضافية مثل المنطق الضبابي لإزالة ضبابية القرار والوصول إلى قرار واضح خال من الغموض.

الكلمات المفتاحية: الجمعيّات الخيريّة، التصنيف، أشجار القرار، خوارزمية C4.5، تقييم النموذج، WEKA

Abstract:

Evaluating and choosing appropriate criteria for decision - making at charities is not a simple process. This process involves defining the needs, evaluating and classifying families, and making the appropriate decisions. Within the period of the crisis between 2012 and 2019, several problems emerged that needed a sound, appropriate and neutral decision at the charities.

The charities set the criteria and mechanism necessary to make a decision. However, error is still present when evaluating the families. The research used the descriptive approach and aimed at identifying and analyzing the criteria, classification of families, and needs according to the algorithm C4. 5. The WEKA tool was used to develop the model for evaluating families, as the decision tree was developed using a vast amount of data. Finally, the decision tree model was evaluated through careful interpretation of the results.

The research indicated the possibility of developing an accurate decision tree; the accuracy degree of the model using the C4.5 algorithm was 0.9691. The model was evaluated through the WEKA tool. The researcher recommended the importance of developing the decision - making process using additional techniques, such as the fuzzy logic.

***Keywords**: Charities, Classification, Decision Tree, C4.5 Algorithm, Model Evaluation, WEKA.*

المقدمة:

لا يوجد حد تستطيع الأعمال الخيريّة الوقوف عنده، فهي متنوِّعة تستطيع أن تشتمل على كافة نواحي وجوانب الحياة، فقد تكون بالتبرعات العينيّة أو النقديّة، أو بمكافحة الأميّة، أو بمكافحة انتشار مرض معين، أو بمساعدة ذوي الاحتياجات الخاصة ورعايتهم، أو ببناء مساكن للفقراء، أو بدعم المشاريع الصناعيّة التي تشغل العاطلين عن العمل، أو بزيارة المرضى والمُسنين، أو بكفالة الأيتام وطلاب العلم، وغيرها العديد من الأعمال الخيريّة التي لا تعد ولا تحصى.والجميل بالأمر أن العمل الخيري يُعطِي صاحبه القدرة على زيادة منسوب إنسانيَّته، فهو يمد صاحبه بطاقة غير عادية.

لذا لا بد من الحاجة لاتخاذ القرارات في ظل البيئة المعقَّدة والمتغيِّرة ديناميكيًا التي تعيشها الجمعيّات في الوقت الحاضر، بالتالي لا بد من استخدام التقنيّات الجديدة التي تعتمد على الذكاء الاصطناعي لبناء نموذج دعم قرار قابل للتكيف مع بيئات مختلفة، إذ يعتمد نجاح أو فشل الجمعيّات على قدرتها في معرفة أي من العائلات تكون محتاجة أو غير محتاجة لتقييم حالات الأشخاص وتقديم المساعدات وغيرها من الأمور التي تحتاج لعمليّة اتخاذ القرار.

وسيتم من خلال هذا البحث التعرُّف إلى آلية تقييم العوائل (المحتاجة وقليلة الاحتياج وغير المحتاجة) ضمن الجمعيّات الخيرية من خلال المعايير التي اعتمدتها الجمعيات الخيرية وطرق اتخاذ القرار، كما سيتم دراسة المعايير وتحليل البيانات الناتجة عن تطبيق هذه المعايير بهدف الوصول إلى آلية يمكن من خلالها تحديد مدى حاجة العوائل إلى المساعدة وبشكل دقيق وذلك من خلال استخدام الأداة WEKA للحصول على شجرة قرار والتنبؤ بتقييم العوائل، وسيتم عرض أهم النتائج والتوصيات التي توصل إليها البحث.

مشكلة الدراسة:

تتمثل مشكلة البحث بأن القرارات التي تحتاج للتقييم وتحديد الاحتياجات (خاصة تقييم العوائل المحتاجة وقليلة الاحتياج و غي المحتاجة) للحصول على الحصة الغذائيّة نتيجة حجم البيانات الكبيرةالتي تحتاج للمعالجة، **نتيجة لذلك ظهرت لدينا التساؤلات التالية:**

- هل يصعب اتخاذ القرار وتحديده بدقة باستخدام الخبرة البشرية؟
- هل الخبرة البشرية موجودة في كثير من الأحيان؟
- هل يتم اعتماد معايير واضحة لاتخاذ القرار؟

- هل تعتبر المعايير كثيرة ومكررة لتقييم العائلات؟

أهمية الدراسة:

تظهر أهمية البحث من استخدام تقنيات حديثة تساعد على التنقيب في البيانات لاستخلاص المعرفة من المعايير التي تم تحديدها لتقييم العوائل، والتي من الصعب استخلاصها باستخدام الطرق التقليديَّة، حيث تساعد المعرفة بالوصول إلى القرار اللازم لتحديد احتياج العوائل، وبالتالي سيكون لهذه المعرفة أهمية ودور كبير في عملية المساعدة في اتخاذ القرار بشكل أفضل وبالوقت المناسب.

كما تنبع أهمية البحث من أهمية عمل الجمعيات الخيرية بما تقدمه من دعم كبير للعائلات المحتاجة وبالتالي إيجاد أسس علمية تساعد على اتخاذ القرار المناسب بالسرعة القصوى من خلال الأداة Weka التي تسهل عملية التنقيب في البيانات إذ تتنبأ بتصنيف العوائل ليتمكنوا من الاستفادة من الدعم الذي تقدمه الجمعيات بشكل كامل دون حدوث ظلم ـأو خطأ.

أهداف الدراسة:

يهدف البحث إلى بناء شجرة قرار لمساعدة متخذ القرار في الجمعيّات الخيريّة على معرفة التقييم المناسب لكل عائلة وفق بيانات خاصة بكل عائلة والتي تم الحصول عليها من جمعية أهل الخير بحلب عبر ملف Excel، وباستخدام الأداة البرمجية WEKA، وتقنية C4.5 المعتمدة على التقسيم والانتزاع بعد أن يتم تدريب الشجرة واختبارها وتقييم مدى دقة نتائجها وصلتها بالبيانات الواقعيّة.

ومن هنا يمكن صياغة التساؤلات الآتية:

أ. ما هي الأساليب العلميّة والتقنيّة المناسبة للتعامل مع البيانات الضخمة وتجهيزها ضمن الجمعيّات لدعم القرارات وتحديد الاحتياجات (خاصة تقييم العوائل المحتاجة وقليلة الاحتياج وغير المحتاجة) ؟

ب. هل يمكن استخدام تقنيّات أشجار القرار لدعم القرارات ضمن الجمعيّات والوصول إلى القرارات المطلوبة ضمن الجوانب المحددة بالدقة المطلوبة وبأسرع وقت ممكن؟

منهجية البحث:

يعتمد البحث على المنهج الوصفي حيث يتم تحليل بيانات العوائل لتقييم حالات العوائل عن طريق بناء شجرة القرار باستخدام WEKA لبناء نموذج لتقييم العوائل عن طريق شجرة القرار المبنية على أساس تقنية C4.5، كما سيتم تقييم النتائج عن طريق مرحلة التدريب بشكل دقيق، إضافة إلى القيام بعملية الاختبار لنموذج التدريب وذلك للتأكد من قدرة وكفاءة ودقة نموذج شجرة القرار الناتجة.

حدود البحث:

- الحدود الزمنية: تغطي الدراسة الفترة من عام (2015) حتى عام (2017) خلال فترة الحرب ضمن سوريا محافظة حلب.
- الحدود المكانية: الجمعية الخيرية ضمن محافظة حلب يطلق عليها (جمعية أهل الخير).
- الحدود الموضوعية: يقتصر موضوع الدراسة الحالية على بناء نموذج أشجار القرار لدعم قرارات الجمعيات الخيرية بحلب ضمن القسم الإغاثي ليتم إعطاء القرار الأدق والأنسب والعادل للعوائل المحتاجة.

الدراسات السابقة:

(Brunello Andrea et al 2019) : بعنوان " معالجة البيانات المتسلسلة والوقتية للحالات المرضية باستخدام أشجار القرار" تم استخدم المنهج التحليلي ضمن الدراسة واعتماد على المتغيرات الطبية لمعالجة حالة مرضية، والتنبؤ من خلال البيانات المتسلسلة والوقتية باستخدام أشجار القرار والوصول إلى القرار والنتائج المطلوبة للمرض بنسبة كبيرة كما توصلت النتائج إلى أنه يمكن معرفة المريض من خلال معرفة (فصيلة الدم والجنس) والسمات الرقمية (العمر والوزن والطول) والبيانات المتسلسلة (تاريخ الأعراض التي تعرض لها، الوصفات الطبية التي أعطيت له، والسلسلة الزمنية تطورت مع مرور الوقت من ضغط دمه، ومعدل ضربات القلب له) ضمن فترات زمنية معينة لكل مريض لدراسة الحالة، قد يكون اتجاه البحث هو تمديد النموذج من أجل التعامل مع صيغ المنطق الزمني، استخدام البحث البرامج الإحصائية بالإضافة إلى رسم شجرة القرار ضمن برنامج Weka

(Vaidehi R, 2016) : بعنوان "مقارنة بين ثلاث تقنيات مختلفة النموذج الخطي العام وأشجار القرار خوارزمية C4.5 والنمذجة القاعدية " استخدمت هذه الدراسة النهج المقترح لتعزيز العملية التنبؤية، وتتناول البيانات الحقيقية المقدمة من آلة UCI. تم جمع مجموعة البيانات من بنك التجزئة البرتغالي، من أيار 2008 إلى حزيران 2013.ترتبط مجموعة البيانات بحملات التسويق المباشر لمؤسسة مصرفية برتغالية، وقد تم استخدام الأداة R وتطبيقها على الدراسة حيث تم تحليل 45211 سجلاً واستخدمت 16 متغيراً تتعلق ببيانات العميل الشخصية وبيانات العميل المصرفية، وبيانات تتعلق بالاتصالات بالعميل، وخلصت الدراسة إلى تحديد 6 متغيرات تكون أفضل في بناء النموذج، وركزت الدراسة على المقارنة بين الأسلوب الأول باستخدام كامل المتغيرات، والأسلوب المقترح باستخدام متغيرات محددة، وخلصت الدراسة إلى أن المتغيرات التي تم تحديدها وهي 6 متغيرات فقط واختبرت بنموذج المصنفات القاعدية وكانت الأفضل

(Abbas S.2015) : بعنوان "مقارنة أساليب شجرة القرار مع أسلوب المجموعات التقريبية اعتمادا على بيانات العملاء المشتركين" .

استخدمت هذه الدراسة منهج التصنيف وقامت بالمقارنة بين أسلوب شجرة القرار المبنيّة بخوارزمية C4.5، وأسلوب المجموعات التقريبيّة Rough Set Theory.اعتمدت الدراسة على 4521 سجلاً من بيانات بنك في برتغالية، تحوي هذه البيانات على 17 متغيرًا تتعلق المتغيرات ببيانات العملاء الشخصيّة، ومعلومات التواصل، واستخدمت أداة Weka بإصدار 11 – 6 – 3 وركزت الدراسة على القواعد المستخرجة وتحليل النموذج دون التركيز على تقييم النموذج، وكانت النتائج تعبر عن أن القواعد التي تم استخراجها من

النموذج تساعد متخذي القرار على أداء القرار بشكل أفضل وأسرع بمنح القروض للأفراد بالإضافة إلى إدارة المخاطر.

(2014، مقداد محمد إبراهيم) بعنوان «دور الجمعيات الأهلية في تخفيض معدلات الفقر في قطاع غزة»

هدفت هذه الدراسة إلى بيان دور جمعية الصلاح الإسلامية في تخفيض معدلات الفقر في قطاع غزة من خلال ما تقدمه من خدمات اقتصادية وتعليمية وصحية للفئات الفقيرة، وذلك خلال الفترة (2004 – 2012م).ولتحقيق أغراض الدراسة قام الباحثون باستخدام المنهج الوصفي والتحليلي الكمي، حيث يعد المنهج الوصفي التحليلي من أكثر المناهج استخداما في دراسة الظواهر الاجتماعية والإنسانية، وهو يناسب الظاهرة موضع الدراسة، فقد ركزت الدراسة على استطلاع رأى المستفيدين من الجمعية من خلال تصميم استبانة لهذا الغرض، حيث استخدم الباحثون طريقة العينة العشوائية، وتم توزيع 550 استبانة على عينة الدراسة، وقد تم استرداد 510 استبانة بنسبة 92.7% واستخدم الباحثون برنامج SPSS لتحليل الاستبانة، في حين يستخدم المنهج الكمي لبناء نموذج قياسي لقياس أثر خدمات الجمعية (الكفالات والصحة والتعليم) في تخفيض معدلات الفقر وذلك باستخدام البرنامج. الإحصائي EViews7، توصي الدراسة جمعية الصلاح الإسلامية بتحسين وتطوير أدائها في مجالات الصحة والتعليم والكفالات؛ لكي تساهم بشكل أكبر في تخفيض معدلات الفقر للفئات المستهدفة، كما تدعو الحكومة لأخذ دورها الفعال في توفير المزيد من الدعم للفئات المهمشة والفقيرة في المجتمع من خلال الوزارات المعنية كوزارة شؤون الاجتماعية وكافة الجهات صاحبة الاختصاص.

(Nwachuckwu , Kabaril) عام 2013 بعنوان "تطوير نموذج دعم قرار باستخدام شجرة القرار والشبكات العصبية لدعم نظام منح القروض البنكية". تم استخدام المنهج الوصفي التحليلي حيث أصبحت اتخاذ القرار في بيئة معقدة ومتغيرة في يومنا هذا بحاجة إلى استخدام تقنيات جديدة في الذكاء الاصطناعي للحصول على نظام دعم قرار ذكي، و قد تم تطبيق النموذج المقترح لمعالجة نظام دعم القرار على طلبات القروض الموجهة للبنوك والوصول إلى القرار المناسب لمنح القروض للزبائن، وتم الاعتمادعلى بيانات البنك في نيجيريا، وتم استخدام ضمن الدراسة لغة برمجية غرضية التوجه (OO - AD) ضمن الأداة Mat lab لتنفيذ الدراسة المقترحة وكان معدل نجاح النموذج 88% ونتائج القرارات جيدة وتساعد متخذي القرار على الوصول إلى القرار بشكل أفضل وأسرع.

(Masud R, Rashedur M 2013) بعنوان "مقارنة أسلوب شجرة القرار وأسلوب بايز في التسويق المصرفي المباشر" استخدمت هذه الدراسة نهج المقارنة وخوارزمية C4.5 لبناء شجرة القرار وأسلوب بايز حيث تم تحليل 4521 سجلاً تم اختيارهم بشكل عشوائي من بيانات UCI المتاحة للجمهور التدريب والاختبار استخدمت هذه الدراسة 15 متغيراً تتعلق المتغيرات ببيانات العملاء الشخصية والبيانات المصرفية للعملاء وبيانات تتعلق بالاتصالات بالعميل.عملت الدراسة أيضا على مقارنة الأسلوبين فقط، باستخدام المقاييس والتصنيفات الصحيحة والخاطئة وذلك من خلال الاعتماد على أداة Weka، كما ركزت الدراسة على استنباط المعرفة الموجودة في البيانات.وقد بينت الدراسة أن نموذج شجرة القرار المبني باستخدام خوارزمية C4.5 هو الأفضل.

(RenukaD.Suryawanshi, D.M.Thakore, 2012) بعنوان "تنفيذ شجرة القرار للتصنيف مع استخدام المنطق الضبابي"

استخدمت الدراسة المنهج الوصفي، تمّ ضمن هذه الدراسة بناء شجرة القرار التي تقوم باستخراج البيانات وتصنيفها بشكل مناسب، ثم تم اختيار المتغيِّرات ذات التأثير الأكبر من غيرها واعتبارها مدخلات المنطق الضبابي للوصول إلى النتائج المطلوبة. بعد ذلك تمّ مقارنة النتائج في حال استخدام شجرة القرار فقط أو شجرة القرار مع المنطق الضبابي لمعرفة الطريقة الأفضل بينهما وذلك باستخدام بيانات سوق الأوراق المالية والحصول على البيانات من مستودع UCI، حيث كانت دقة النموذج 98% وهي دقة عالية.

(Jože BENČINA, 2011) بعنوان "أشجار القرار الضبابيّة لصنع القرار في القطاع العام"

استخدمت الدراسة منهج التصنيف والتحليل في استخدام أشجار القرار والمنطق الضبابي» لدراسة حالة الاستثمار وتقييم المشاريع ضمن قسم البلدية في سلوفينيا، بالإضافة لتقييم كفاءة وفعاليّة المؤسسات العامة وذلك بإيجاد حلول لها خصائص معدّلة خصيصًا للاستخدام في القطاع العام، ويهدف البحث إلى جعل تطبيقات التقييم تقدم تقييما دقيقا وبشكل أفضل، وتكون سهلة الوصول والاستخدام، ويمكن تطبيق النموذج للتقييم في كافة المجالات ضمن القطاعين العام وغير العام، وتم استخدام مكتبة برمجة خاصة لبناء النموذج وكانت نتائج البحث دقيقة وجيدة لمساعدة متخذي القرار للحصول على قرار دقيق وواضح.

تم عرض مجموعة من الدراسات السابقة تعبر عن استخدام تقنية أشجار القرار في بناء نماذج لصنع القرار، وقد ظهر ضمن النماذج أن أسلوب شجرة القرار من الأساليب المفيدة والجيدة للوصول إلى قرار مناسب وواضح بأسرع وقت وأفضل دقة.

ما يميز الدراسة الحالية عن الدراسات السابقة:

ما يُميِّز هذه الدراسة عن الدراسات السابقة أنه سوف تُطبَّق على الجمعيّات الخيريّة، وسوف تقوم بتوصيف المعايير التي وضعتها المنظّمات العالميّة، ودراسة ديناميكيّة هذه المعايير ومعالجتها من خلال مراعاة أسس التقييم والتي على أساسها يتم اتخاذ القرار، بالإضافة إلى العوامل المؤثرة التي وضعتها الجمعيّات لإجراء عمليّة التصنيف لحالات العائلات للوصول إلى المتغيِّرات ذات التأثير الأكبر على عمليّات التصنيف وذلك باستخدام الأداة WEKA.

أولاً: الإطار النظري:

مفهوم الجمعيّات الخيريّة:

لا يمكن لأي دولة بمؤسساتها المختلفة مهما بلغت من درجات في التقدم والازدهار والنمو الاقتصادي والرفاه المعيشي من تحقيق جميع الاحتياجات اللازمة لاستمرار عجلة الحياة في مجتمعها، إذ تعد حاجات الإنسان متعددة ومتنوعة وغير محددة مقارنة بموارده التي تظل بمستوى محدود.

إن أهميّة الجمعيّات الخيريّة في مختلف دول العالم وخاصة في البلدان النامية تنبع من أنها تسد الثغرات وتلبي احتياجات السكان خاصة في مجال الرعاية الاجتماعية، إذ تقوم هذه الجمعيّات الخيريّة بتلبية حاجات السكان الأكثر فقرًا من خلال تقديم الخدمات ومختلف وسائل العيش ومقومات الحياة.(مقداد ، 2014 ص2)

ولا يمكن للجمعيّات الخيريّة القيام بعملها الإنساني على أكمل وجه وتحقيق هدفها المتمثل في إرضاء الفقراء وسد احتياجاتهم دون وجود تخطيط في كافة جوانب عملها، وتحوّل عملها من العمل العشوائي إلى العمل المؤسساتي المخطط الهادف. (هلال علي 2012 ص5)

فقد كان عمل الجمعيّات الخيريّة في مدينة حلب قبل الأزمة محدودًا يقتصر على الحصول على الأموال من التجار، وبعض الأشخاص الميسورين في المجتمع لتوزيعها على الأشخاص الفقراء، لكن مع ظهور الأزمة توَجّب على المنظمات الدوليّة التدخل لسد حاجات السكان المتضررين بسبب الحرب، وبما أن هذه المنظمات كانت جهات خارجيّة أجنبيّة ليس لها دراية بالبلد؛ عمدت المنظمات إلى توقيع اتفاقيات مع الجمعيّات الخيريّة المحليّة التي أصبحت حلقة الوصل بين المنظمات الدوليّة من جهة وبين السكان المحتاجين من جهة أخرى، فأصبحت تلك الجمعيّات الخيريّة المحليّة تحصل على المساعدات بأنواعها المختلفة لتقديمها إلى السكان المتضررين من الحرب، ومن هنا ظهرت المشكلة؛ هل استطاعت الجمعيّات الخيرية تلبيه حاجات السكان والقيام بواجبها المطلوب بالشكل الصحيح؟ .(عبدربه مجدي 2013 ص 24)

مفهوم تقييم العوائل ضمن الجمعيّات الخيريّة:

تحوَّلت الجمعيّات الخيريّة في مدينة حلب في ظل الأزمة من الاعتماد على الممولين المحليين إلى الممولين الخارجيين المتمثلين بالمنظمات الدوليّة، ومن أبرز الممولين:

برنامج الأغذية العالمي (PFW) :

حيث يقوم برنامج الأغذية العالمي بتوفير الحصص الغذائيّة، وإرسالها إلى الجمعيّات الخيريّة لتقوم بدورها بتوزيعها على مستحقيها.وقد اعتمدت الجمعيّات الخيريّة على برنامج Excel فقد قامت كل جمعية خيرية بإنشاء ملف Excel يتم من خلاله تقييم العائلة؛ حيث تقوم الجمعيّات الخيريّة بجمع بيانات معايير معينة تم تحديدها على ملف Excel من قبل منظمة WFP، بعد عملية جمع المعايير يتم اتخاذ القرار وتحديد حالة العائلة؛ هل هي من العوائل كبيرة الاحتياج أو متوسطة الاحتياج أو قليلة الاحتياج.**وتتمثل المعايير ضمن المنظّمات (حسب المعايير من أداة المنظمة WFP) في الشكل التالي:**

- تفاصيل عن الأسرة: (الاسم، رقم الهاتف، أعمار أفراد الأسرة، مسؤول الأسرة هل هو ذكر أم أنثى لأنه في حال كان المسؤول أنثى فتختلف عمليّة تقييم الأسرة، الوضع الصحي لرب الأسرة).
- تصنيف الأسرة: تصنّف الأسرة حسب عدد الأفراد.
- نسبة الاتكاليّة: تتوضح نسبة الاتكاليّة حسب عمر الأفراد ضمن الأسرة.
- حالة معيل الأسرة: تأخذ حالة معيل الأسرة القيم (هل هو طفل، مسن...إلخ)
- حالة الهشاشة: وتتمثل حالة الهشاشة بمعرفة هل معيل الأسرة معاق، مصاب، المسؤول زوجة وهي حامل، عدد الأفراد المعاقين ضمن الأسرة والتي تتراوح أعمارهم من 16 لـ 59.
- وضع الإقامة: (مقيم، مُهجّر، مُهجّر قديماً، مُهجّر حديثاً، مُهجّر بشكل متكرر، مُهجّر وعائد للمنزل).
- وضع المأوى: (حالة عامة، البيت على الهيكل، منازلهم مدمرة، مستأجر).
- مصادر الدخل للأسرة: (زراعيّة في حال يملك أراضي زراعيّة، غير زراعيّة (موظف، محامي، طبيب)، مصادر أخرى، لا يوجد دخل في حال كان الدخل أقل من 25000 ل.س).

بعد ذلك يتم ترجيح المعايير بنقاط معينة ليتم جمع تلك النقاط وتحديد حالة العائلة هل هي عائلة كبيرة الاحتياج أو متوسطة الاحتياج أو عديمة الاحتياج.ولكن بعد تحليل البيانات ودراستها وتقييمها، وقد تكون البيانات السابقة كثيرة وقد لا تكون متوفرة دوماً، وبالتالي من الصعب تقييمها، نلاحظ أنه تم اختصار البيانات إلى أقل حد وبالتالي سهلت الحصول على هذه البيانات وتحليلها، والسؤال هو: هل ستفي هذه البيانات بالغرض في الجواب على السؤال من الدراسة التطبيقية؟ و سوف يعتمد البحث على العناصر التالية لتقييم العوائل: (عدد الذكور، عدد الإناث، تصنيف حجم العائلة، هل تحتوي العائلة على أطفال أو مسنين، تصنيف نسبة الاتكاليّة، تصنيف حالة رب الأسرة، تصنيف فئة الهشاشة بالأسرة، تصنيف التهجير للأسرة، تصنيف مأوى الأسرة، نوع الدخل).

ويركز هذا البحث بشكل أساسي على بناء شجرة القرارات وبشكل خاص تقنية C4.5 وتم الاعتماد على الأداة WEKA؛ لأنها أداة مفتوحة المصدر، وذات كفاءة وقدرات تحليليّة عالية، ويمكن تحديثها من خلال الموقع الخاص بالأداة بشكل مستمر.

الأداة Weka:

عبارة عن حزمة برمجية مفتوحة المصدر تحتوي على مجموعة من الخوارزميات يمكن تطبيقها بسهولة على مجموعة من البيانات إما بشكل مباشر من خلال واجهة برنامج Weka، أو من خلال استدعائهاعن طريق (Java Code) باستخدام Classes الخاصة بها، وذلك من خلال تحميل المكتبة الخاصة Weka، وأولى خطوات التحليل باستخدام Weka هي استخدام بيانات ذات صيغة قابلة للفهم من قبل Weka أحد هذه الصيغ هي صيغة arff.

يعتمد هذا البحث على استخدام الأداة Weka لأنها أداة مفتوحة المصدر، وذات كفاءة وقدرات تحليلية عالية، ويمكن تحديثها من خلال الموقع الخاص بالأداة بشكل مستمر.بدأ تطوير هذه الأداة منذ عام 1997 والنسخة الأخيرة تعتمد بشكل أساسي على لغة الجافا، يمكن استخدام هذه الأداة في مجالي البحث العلمي والتعليم.(Aksenova S Svetlana, 2004, p2)

تقييم نموذج شجرة القرار:

التقييم هو المفتاح الأساسي الذي يمكن من خلاله إحداث تقدم حقيقي بعد عملية معالجة البيانات، ولكن يجب التركيز على تقييم النتائج التي حصلنا عليها، ولا تُعد عملية التقييم بسيطة كما قد تبدو للوهلة الأولى، لقد اعتمدت بعض الأبحاث في التقييم على نتائج التدريب فقط، وهذا غير دقيق، نحن بحاجة للتأكد من أن نتائج الأداء التي حصل الباحث عليها في عملية التدريب هي نتائج دقيقة ويمكن الاعتماد عليها وذلك من خلال القيام بعملية الاختبار، حيث تتم عملية الاختبار باستخدام بيانات مستقلة عن البيانات التي تم استخدامها للتدريب، ونقوم بدراسة دقة النتائج التي حصل عليها الباحث نتيجة الاختبار، ونقارنها مع النتائج التي حصل عليها في مرحلة التدريب، فإن كانت القيم التي تعبر عن دقة النتائج في كلا مرحلتي التدريب والاختبار متقاربة؛ فهذا مؤشر على أن نتائج التدريب جيدة ويمكن الاعتماد عليها، وبهذا الشكل يكون قد قمنا بعملية التقييم على أكمل وجه.بعد تحديد معدل الخطأ للنموذج الذي تم استخدامه يمكن إعادة بيانات الاختبار إلى بيانات التدريب (دمج مجموعتي بيانات التدريب والاختبار مع بعضهما) وذلك بهدف زيادة حجم بيانات التدريب الجديدة ثم القيام بإعادة توليد النموذج اعتمادًا على مجموعة التدريب الأخيرة، وبهذا الشكل تصبح البيانات المستخدمة للتدريب جاهزة، وستكون مقاييس تقييم الأداء أفضل من السابق.(ساكت، عداس 2015 ص8)

مقاييس تقييم أداء التصنيف:

يوجد العديد من المقاييس التي يمكن استخدامها من أجل تقييم أداء التصنيف، منها: إحصائية Kappa، ومعدل الخطأ المطلق، وجذر متوسط مربع الخطأ، ودقة التصنيف Precision، ونسبة التصنيف الصحيح الإيجابي (TP Rate)، و Recall (الاستدعاء)، و F - Measure، و Roc Area.سوف نقوم أولاً بتوضيح حالات التصنيف، وتمثيل مصفوفة التشويش والتي سوف تستخدم في حساب قيمة مقاييس أداء التصنيف. (WITTENN I., et al , 2011 p5)

تبدأ عملية التقييم من خلال مصفوفة التشويش حيث تبين هذه المصفوفة عدد السجلات التي تم التنبؤ بها بشكل صحيح بالنسبة لكل صنف من الأصناف.فبفرض لدينا نموذج يقوم بالتنبؤ بصنفين **a=Yes** و **b= No** **فإن مصفوفة التشويش تبين الحالات المختلفة للصنفين كما يلي:**

1. الصحيحة الإيجابية (TP) : إذ تكون الحالة صحيحة إيجابية عندما تكون قيمتها الفعلية إيجابية وتم التنبؤ بها أو تصنيفها على أنها إيجابية.
2. الصحيحة الإيجابية (TN) : حيث تكون الحالة صحيحة سلبية عندما تكون قيمتها الفعلية سلبية وتم التنبؤ بها أو تصنيفها على أنها سلبية.
3. الصحيحة الإيجابية (FP) : حيث تكون الحالة صحيحة إيجابية عندما تكون قيمتها الفعلية سلبية وتم التنبؤ بها أو تصنيفها على أنها إيجابية.
4. الصحيحة الإيجابية (FN) : حيث تكون الحالة صحيحة إيجابية عندما تكون قيمتها الفعلية إيجابية وتم التنبؤ بها أو تصنيفها على أنها سلبية.

- Kappa: تعبر إحصائية Kappa عن نسبة النجاح المثاليّة للتصنيف، إذ تقيس التوافق بين التصنيف والقيم الفعلية بعد تطبيع أو طرح التوافق العشوائي الناتج من الصدفة، وتتراوح قيمة هذا المقياس بين الصفر والواحد، فعندما تأخذ قيمة 1 يكون التوافق مثاليا تاما بين التصنيف والقيم الفعلية، أما عندما تأخذ القيمة 0 فيكون التوافق عشوائيا ناتجا من الصدفة.وكلما زادت قيمة هذا لملمقياس زادت دقة كفاءة نموذج التصنيف، فإذا كانت قيمة Kappa ضمن المجال [59. – 40.] عندئذ درجة النجاح في التصنيف معتدلة أما إذا كانت قيمة Kappa ضمن المجال [79. – 60.] فعندئذ درجة النجاح في التصنيف كبيرة، وأخيرًا إذا كانت قيمة Kappa ضمن المجال [1.00 – 80.] عندئذ درجة النجاح في التصنيف ممتاز.
- MAE : يعبر MAE عن معدل الخطأ المطلق بين القيم الفعليّة والقيم المتنبأ بها، وكلما انخفضت قيمة هذا المقياس زادت دقة وكفاءة نموذج التصنيف.
- RMSE : يعبر RMSE عن جذر معدل مربع الخطأ النسبي بين القيم الفعلية والقيم المتنبأ بها، وكلما انخفضت قيمة هذا المقياس زادت دقة وكفاءة نموذج التصنيف، ويعتبر جذر معدل مربع الخطأ المطلق وكلما انخفضت قيمة هذا المقياس RMSE زادت دقة وكفاءة نموذج التصنيف.
- RAE : يعتبر الخطأ المطلق النسبي RAE أكثر ملاءمة من معدل الخطأ المطلق، وكلما انخفضت قيمة هذا المقياس RAE زادت دقة وكفاءة نموذج التصنيف.
- Recall : مقياس التقييم Recall (الاستدعاء) TP Rate= ويعبر عن النواحي التي صنفت على أنها إيجابية صحيحة وهي في الواقع إيجابية صحيحة.
- Precision : يسمى بمعامل التحقيق ويحدد هذا المقياس نسبة السجلات التي يتبين أنها إيجابية والتي صرح المصنف بأنها إيجابية.
- F_Measure : مقياس التقييم F_Measure يجمع بين مقياسي Precision و Recall
- ROC Area : يعبر مقياس التقييم ROC Area عن قدرة النموذج على تصنيف بيانات الاختبار بشكل صحيح.وهو يعتبر من أهم المقاييس، وأقل قيمة يأخذها صفر وأكبر قيمة يأخذها واحد، ويكون التصنيف أفضل كلما اقتربت القيمة من الواحد.عندما تصل قيمة هذا المقياس إلى 5. فإنه سوف يكون في منطقة عدم التمييز وذلك سوف يؤثر بشكل سلبي على تعليم نموذج التصنيف (ساكت، عداس 2015، ص10، 11)

ثانياً: الدراسة التطبيقية:

يتم تقييم قرار احتياج العوائل من خلال المعايير التي تم

اختيارها من منظمة WFP لإمكانية معرفة تأثير كل المعايير على القرار واحتياج كل عائلة، ولكن يوجد صعوبة أثناء التقييم ضمن ملف Excel لذا اضطرت الجمعيّات البحث عن سبل جديدة أخرى تسهِّل عملية التقييم وتعتمد على تحاليل متقدمة.

إن البيانات التي سيتم استخدامها في الحالة التطبيقيّة تم أخذها من جمعيّة أهل الخير قسم المعونات الغذائيّة، وتتضمن 2173 سجلا، وقام الباحث بتهيئة البيانات ووضعها بشكل يناسب عملية التصنيف، كما تم تقسيم هذه البيانات إلى مجموعتين؛ مجموعة أولى لتدريب وبناء النموذج، ومجموعة ثانية لاختبار النموذج.تم بناء نموذج أشجار القرار باستخدام خوارزمية C4.5 بالاعتماد على المجموعة الأولى من البيانات، كما تم اختبار النموذج الناتج عن مرحلة التدريب على المجموعة الثانية من البيانات (بيانات الاختبار) وذلك باستخدام برنامج WEKA.

تتكون قاعدة البيانات من 11 متغيرا، 10 متغيرات مستقلة وهي (عدد الذكور، عدد الإناث، تصنيف حجم العائلة (صغيرة، كبيرة، متوسطة)، هل تحتوي العائلة على أطفال أو مسنين (طفل، مسن، لا يوجد)، تصنيف نسبة الاتكاليّة (مرتفعة، متوسطة، منخفضة)، تصنيف حالة رب الأسرة (رجل، امرأة، طفل، معاق، مسن، لا يوجد)، تصنيف فئة الهشاشة بالأسرة (مرتفعة، متوسطة، منخفضة)، تصنيف التهجير للأسرة (تهجير قديم، تهجير متعدد، عائد)، تصنيف مأوى الأسرة (أبنية عامة، مستأجر، مدمر جزئيًّا)، نوع الدخل (قليل، متوسط، أخرى، لا يوجد)، بالإضافة إلى متغير تابع هو تحديد نوع الاحتياج، في حال كانت الحالة A تعبر عن العائلة المحتاجة، الحالة B العائلة متوسطة الاحتياج، الحالة C العائلة قليلة الاحتياج، و سوف نوضح المتغيرات المستخدمة ضمن الجدول رقم (1):

جدول شرح المتغيرات المستخدمة في الدراسة رقم (1)

الشرح	النوع	المتغير
عدد الذكور ضمن الأسرة.	رقمي	Male
عدد الإناث ضمن الأسرة.	رقمي	Female
تصنيف حجم العائلة (صغيرة، كبيرة، متوسطة).	فئوي	classification1
هل تحتوي العائلة على أطفال أو مسنين (طفل، مسن، لا يوجد).	فئوي	Is Child Or Aged?
تصنيف نسبة الاتكاليّة (مرتفعة، متوسطة، منخفضة).	فئوي	Classification Dependency
تصنيف حالة رب الأسرة (رجل، امرأة، طفل، معاق، مسن، لا يوجد).	فئوي	Classification Status
تصنيف فئة الهشاشة بالأسرة (مرتفعة، متوسطة، منخفضة).	فئوي	Classification Fragility
تصنيف التهجير للأسرة (تهجير قديم، تهجير متعدد، عائد).	فئوي	Classification Displacement
تصنيف مأوى الأسرة (أبنية عامة، مستأجر، مدمر جزئيًّا).	فئوي	Classification House
نوع الدخل (قليل، متوسط، أخرى، لا يوجد).	فئوي	Salary
قرار احتياج الأسرة حيث المجموعة A وهي الأسر التي تمتلك درجة عالية من الاحتياج. المجموعة B وهي الأسر التي تمتلك درجة متوسطة الاحتياج. المجموعة C وهي الأسر التي تمتلك درجة عديمة الاحتياج.	فئوي	Basic

بناء شجرة القرار وتقييمها:

عن طريق برنامج الـ WEKA قام الباحث بتحضير البيانات المراد العمل عليها، و واختار نوع التصنيف شجرة القرار من نوع J48، ويعود سبب اختيار هذا النوع من أشجار القرار لدقة نتائجه مقارنة بالأنواع الأخرى من أشجار القرار وذلك حسب ما توصلت إليه الدارسات السابقة.

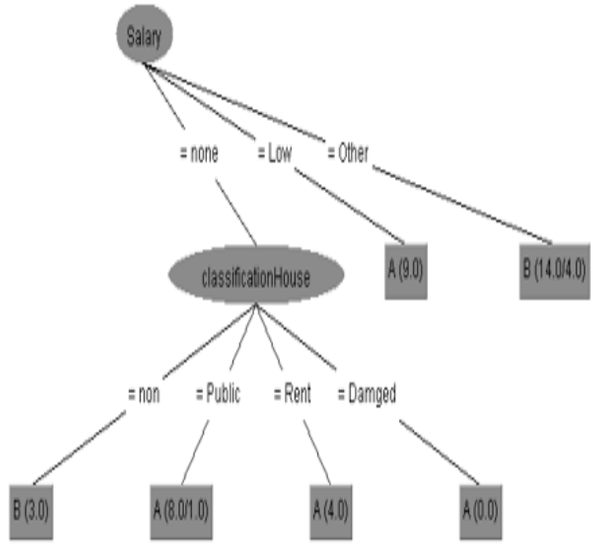
سوف تقوم الدراسة العمليّة على مرحلتين هما: مرحلة التدريب، ومرحلة الاختبار، مع التقييم الدقيق لنتائج كل مرحلة.

1. مرحلة التدريب:

تضمنت مجموعة البيانات 2173 سجلا، تم استخدام في مرحلة التدريب 2149 سجلا لبناء نموذج شجرة القرار وبعد بناء النموذج حصل الباحث على شجرة القرار التالية:

2. شجرة القرار:

تساعد هذه الشجرة على اتخاذ القرار بالنسبة لأي عائلة جديدة تقدمت بطلب الحصول على مساعدة حيث يتم أولاً السؤال عن الراتب للعائلة وبناء على الجواب يتم التقدم والمتابعة بحسب بنية الشبكة الموضحة بالشكل التالي:



شكل رقم (1)

جزء من شجرة القرار باستخدام تقنية J48

تقييم نتائج مرحلة التدريب:

لتقييم شجرة القرار التي تم الحصول عليها في مرحلة التدريب لا بد من دراسة ملخص عملية التقييم المعطاة بالجدول رقم (2)

دراسة ملخص عملية التقييم المعطاة بالجدول رقم (2)

ملخص عملية التقييم لمجوعة بيانات الاختبار		
	عدد السجلات	القيمة
نسبة السجلات المصنفة بشكل صحيح	2149	98.8955%
نسبة السجلات المصنفة بشكل خاطئ	24	1.1045%
العدد الكلي للسجلات	2173	
معامل Kappa		.9691
معدل الخطأ المطلق		.0132
الجذر التربيعي لمعدل مربع الخطأ		.0812
الخطأ المطلق النسبي		5.4408٪

ملخص عملية التقييم لمجوعة بيانات الاختبار		
	عدد السجلات	القيمة
الجذر التربيعي للخطأ النسبي		23.3389٪

يبين ملخص عملية تقييم مجموعة التدريب أن عدد حالات التصنيف الصحيح 2149 حالة ونسبتها المئوية 98.8955، في حين أنّ عدد حالات التصنيف الخاطئ 24 حالة ونسبتها المئوية 1.1045، العدد الكلي للحالات 2173 حالة، وهي نتائج جيدة جدًا إذ كانت قيمة إحصائية Kappa إلى 9691.، وتعبر عن نسبة نجاح النموذج، فهي تقيس التوافق بين القيم المتوقعة والقيم الفعليّة وتعتبر النسبة كبيرة، و متوسط الخطأ المطلق بين القيم الفعليّة والقيم المتنبأ بها يساوي 0132.، وقد كان متوسط مربع الخطأ بين القيم الفعليّة والقيم المتنبأ بها يساوي 0812.، و النسبة المئوية للخطأ المطلق النسبي بين القيم الفعليّة والقيم المتنبأ بها يساوي 5.4408 %، أما النسبة المئوية للجذر التربيعي للخطأ النسبي للقيم الفعليّة والقيم المتنبأ بها فيساوي 23.3389 %.

يوضح الجدول رقم (5) البيانات التفصيليّة المعبرة عن دقة التصنيف من أجل كل قيمة من قيم دالة الهدف

جدول تفاصيل دقة التصنيف لمجموعة التدريب رقم (3)

	TP Rate	FP Rate	Precision	Recall	F - Measure	PRC Area	Class
	.998	.045	.988	.998	.993	.998	B
	.938	.001	.996	.938	.966	.977	C
	.976	.001	.990	.976	.983	.988	A
Weighted Avg	.989	.035	.989	.989	.989	.995	

يتضمن الجدول رقم (3) الأعمدة التالية:

- العمود TP Rate: يعبر عن نسبة التصنيف الصحيح الإيجابي لكل من حالات العوائل كبيرة الاحتياج ومتوسطة الاحتياج وقليلة الاحتياج، فمثلاً القيمة 998. تعبر عن نسبة التصنيف الصحيح في حالة الأسرة متوسطة الاحتياج، ونلاحظ أن قيم نسبة التصنيف الصحيح الإيجابي مرتفعة في حالة كانت العائلة متوسطة الاحتياج، في حين تعبر القيمة 938. عن نسبة التصنيف الصحيح في حال كانت العائلة قليلة الاحتياج، وتعبر القيمة 976. عن نسبة التصنيف الصحيح في حال كانت العائلة كبيرة الاحتياج، والصف الأخير يعبر عن متوسط نسبة التصنيف الصحيح الإيجابي بين الحالات الثلاثة (A,B,C) ويساوي 989. النسبة المئوية للتصنيف الصحيح.

- العمود FP Rate: يعبر عن نسبة التصنيف الخاطئ الإيجابي لكل من حالات العوائل كبيرة الاحتياج (A) ومتوسطة الاحتياج (B) وقليلة الاحتياج (C)، فمثلاً القيمة 045. تعبر عن نسبة التصنيف الخاطئ في حالة الأسرة متوسطة الاحتياج، ونلاحظ أن قيم نسبة التصنيف الخاطئ الإيجابي منخفضة في حالة كانت العائلة محتاجة، في حين تعبر القيمة 001. عن نسبة التصنيف الخاطئ في حال كانت العائلة قليلة الاحتياج، وعالية الاحتياج نفس نسبة التصنيف الخاطئ، والصف الأخير يعبر عن متوسط نسبة التصنيف الخاطئ الإيجابي بين الحالات الثلاثة (A،B،C) ويساوي 035. النسبة المئوية للتصنيف الخاطئ.

- العمود Precision: يعبر عن نسبة دقة التصنيف لكل من حالات العوائل، فمثلا 988. تعبر عن نسبة دقة التصنيف في حال الأسرة متوسطة الاحتياج، وفي حين تعبر القيمة 996. عن نسبة دقة التصنيف في حالة الأسرة قليلة الاحتياج، في حين تعبر القيمة 990. عن نسبة دقة التصنيف في حالة الأسرة كبيرة الاحتياج، والصف الأخير يعبر عن متوسط نسبة دقة التصنيف ويساوي 989. النسبة المئوية بين حالات الأسرة الثلاثة.

- الأعمدة Roc Area, F - Measure, Racall: عبارة عن مقاييس مختلفة لتقييم النواحي الإيجابية الخاطئة مقابل المبادلات السلبية الخاطئة ويعتبر المقياس Roc Area الأكثر أهمية، وأقل قيمة يأخذها صفرا، وأكبر قيمة واحد، ويكون التصنيف أفضل كلما اقتربت قيمة هذا المقياس من الواحد، و إذا قلت قيمة هذا المقياس

عن 5. فإن ذلك سوف يؤثر بشكل سلبي على تعليم نموذج التصنيف، والصف الأخير من الأعمدة يعبِّر عن متوسط المقاييس بين حالات الأسرة الثلاثة.

نلاحظ من مصفوفة التشويش لنموذج من الجدول رقم (4) أن في حالة الأسرة كبيرة الاحتياج كان عدد الحالات المصنفة بشكل خاطئ 3 (FP)، وفي حالة الأسرة متوسطة الاحتياج كان عدد الحالات المصنفة بشكل خاطئ 16 (FP)، وفي حالة الأسرة قليل الاحتياج كان عدد الحالات المصنفة بشكل خاطئ 5 (FP)، إذ كانت نسبة الخاطئ الكلي هي (24) سجلا.

جدول رقم (4)

مصفوفة التشويش

		القيم المتنبأ بها		
		A	B	C
القيم الفعلية	A	1699	1	2
	B	16	243	0
	C	5	0	207

مرحلة الاختبار:

بعد انتهاء مرحلة التدريب يتم حفظ النموذج لاستخدامه في مرحلة الاختبار، إذ يتم تحميل نموذج التدريب، ومن ضمن خيارات الاختبار يتم اختيار Supplied test set ونختار الملف الذي يحتوي على بيانات الاختبار والمؤلفة 241 سجلا ثم نقوم بتقييم النموذج فنحصل على ملخص البيانات المعطى بالجدول رقم (5)

الجدول رقم (5)

ملخص عملية تقييم لمجموعة بيانات الاختبار

ملخص عملية التقييم لمجموعة بيانات الاختبار		
	عدد السجلات	القيمة
نسبة السجلات المصنفة بشكل صحيح	234	97%
نسبة السجلات المصنفة بشكل خاطئ	7	3%
العدد الكلي للسجلات	241	
معامل Kappa		.9335
معدل الخطأ المطلق		.0308
الجذر التربيعي لمعدل مربع الخطأ		.1241
الخطأ المطلق النسبي		1.277%
الجذر التربيعي للخطأ النسبي		32.1535%

يبين ملخص عملية تقييم مجموعة الاختبار أن عدد حالات التصنيف الصحيح 234 ونسبتها المئوية 97، في حين أنّ عدد حالات التصنيف الخاطئ 7 ونسبتها المئوية 3، العدد الكلي للحالات 241، وهي نتائج جيدة جدًا إذ كانت قيمة إحصائية Kappa إلى 9335.، وانخفاض قيمة متوسط الخطأ المطلق بين القيم الفعليّة والقيم المتنبأ بها ليصبح 0308.، وانخفاض قيمة جذر متوسط مربع الخطأ بين القيم الفعليّة والقيم المتنبأ بها لتصبح 1241.، وانخفاض النسبة المئوية للخطأ المطلق النسبي بين القيم الفعليّة والقيم المتنبأ بها لتصبح 10.277% وانخفاض النسبة المئوية للجذر التربيعي للخطأ النسبي لتصبح 32.1535%.

جدول رقم (6)

تفاصيل دقة التصنيف لمجموعة الاختبار

	TP Rate	FP Rate	Precision	Recall	F - Measure	PRC Area	Class
	.994	.086	.966	.994	.980	.995	B
	.880	.000	1.000	.880	.936	.961	C
	1.000	.005	.952	1.000	.976	.983	A
Weighted Avg	.971	.061	.972	.971	.970	.987	

بمقارنة الجدول رقم (6) مع جدول تفاصيل التصنيف لعملية التدريب رقم (3) نجد شبه تطابق النتائج ما بين الجدولين وهذا يدل على دقة وكفاءة نموذج التصنيف الذي تم تطبيقه وهو ما تؤكده أيضًا نتائج مصفوفة التشويش لمجموعة الاختبار.

جدول رقم (7)

مصفوفة التشويش

		القيم المتنبئ بها		
		A	B	C
القيم الفعلية	A	170	0	1
	B	6	44	0
	C	0	0	20

نلاحظ من مصفوفة التشويش لنموذج من الجدول رقم (7) أنه في حالة الأسرة كبيرة الاحتياج كان عدد الحالات المصنفة بشكل خاطئ 1 (FP)، و في حالة الأسرة متوسطة الاحتياج كان عدد الحالات المصنفة بشكل خاطئ 6 (FP)، و في حالة الأسرة قليلة الاحتياج كان عدد الحالات المصنفة بشكل خاطئ 0 (FP)، أي أن عدد السجلات الكلي المصنفة بشكل خاطئ في مرحلة الاختبار هي (7) سجلات.

تم التأكد بعد تقارب نتائج كل من مرحلتي التدريب والاختبار وتم الحصول على شجرة قرار مع نتائج أكثر دقة.

الاستنتاجات:

1. قدرة شجرة القرار على دعم قرارات الجمعيات الخيرية للأسرة المدخلة وتحديد احتياجها؛ هل هي من صنف الأسر المحتاجة أو الأسر متوسطة الاحتياج أو الأسرقليلة الاحتياج.

2. قدرة الأداة WEKA على بناء شجرة القرارات مدعومة بالعوامل التي توضح دقة النموذج المقترح لدعم قرارات الجمعيات الخيرية.

3. سهولة استخدام شجرة القرار وسرعة الحصول على النتائج والقدرة على التعامل مع البيانات الكميّة والنوعيّة لدعم قرارات الجمعيات الخيرية.

4. بالرجوع إلى نتائج خوارزمية C4.5 تبين أن تم التقارب بين نتائج التدريب والاختبار و هذا يدل على قوة الدراسة.

5. من خلال النموذج الذي تم بناؤه عبر خوارزمية C4.5 تم معرفة أي المعايير المؤثرة على عملية اتخاذ القرار وهي (نوع الدخل، معايير حجم العائلة، تصنيف مأوى الأسرة، تصنيف حالة رب الأسرة، تصنيف فئة الهشاشة بالأسرة).

6. أظهرت النتائج أن عدد حالات التصنيف الصحيحة في مجموعة التدريب ونسبتها ( ) ومجموعة الاختبار ونسبتها ( ) وهي نسبة جيدة وأن عدد حالات التصنيف الخاطئ في مرحلة التدريب ونسبتها ( ) وفي مرحلة الاختبار ونسبتها ( ) وكانت النتائج جيدة ومقبولة.(اكتب النسب)

7. أظهر مقياس TP Rate باستخدام خوارزمية C4.5 أن احتياج العوائل ضمن القسم الإغاثي في أحد الجمعيات الخيرية بلغ () وهي نسبة جيدة.

8. أظهر تحليل مصفوفة التشويش أن نموذج أشجار القرار يمكن نوعا ما الاعتماد عليه في عملية اتخاذ القرار بعد استنتاج قواعد الاستدلال من الشجرة وبناء نموذج المنطق الضبابي ليتم الوصول إلى القرار بنسب دقيقة وواضحة والمعايير ذات التأثير الكبير على القرار.

التوصيات:

1. نوصي باستخدام أشجار القرار والاستفادة منها في تصنيف احتياج الأسر لدعم القرار والحصول على قرار بشكل أفضل.

2. استخدام تقنيات أخرى كالمنطق الضبابي للوصول إلى دعم قرار بشكل سليم.

3. يمكن تطبيق أدوات أخرى وخوارزميات تصنيف أخرى واستخدام طرق نمذجة حديثة ضمن الذكاء الصنعي تساعد في عملية اتخاذ القرار.

4. يمكن تطبيق خوارزميات تصنيف أخرى مثل ID3, Bayesian وذلك لمعرفة جودة تصنيف أي خوارزمية بشكل أفضل للوصول إلى النتائج المطلوبة ومعرفة أي معايير ذات التأثير على القرار بشكل أكبر ضمن كل خوارزمية.

5. يمكن الاعتماد على النموذج وتطبيقه ضمن الجمعيات الخيرية بحلب أو أي محافظة أخرى ضمن القسم الإغاثي لمساعدة متخذي القرار على اتخاذ قرار احتياج العوائل ضمن القسم الإغاثي في الجمعيات الخيرية

6. يمكن تطوير النظام بتعديله إلى إضافة أو تعديل معايير يمكن من خلاله تطبيق النموذج على الأقسام الأخرى لمساعدتهم في عملية التقييم واتخاذ القرار.

7. تطوير آلية العمل إلى الطرق الحديثة ضمن الجمعيات واستخدام الأنظمة الخبيرة مثل النموذج المطبق.

المصادر والمراجع:

أولاً: المراجع باللغة العربية:

- ساكت غسان، عداس ضحى 2015» استخدام شجرة القرار لدعم قرارات منح القروض «، مجلة بحوث جامعة حلب، سلسلة العلوم الاقتصادية العدد 14
- عبدربه مجدي محمد مصطفى2013» التحديات الإدارية التي تواجه الجمعيّات الخيريّة وسبل مواجهتها دراسة تطبيقيّة على عينة من الجمعيّات الخيريّة في سلطنة عمُان»، جامعة السلطان قابوس، مجلة الآداب والعلوم الاجتماعية، سلطنة عمان.
- مقداد محمد إبراهيم 2014» دور الجمعيّات الأهلية الفلسطينيّة في تخفيض معدلات الفقر في قطاع غزة «، الجامعة الإسلامية، غزة.
- هلال علي هلال زينب 2012 «دور الجمعيّات الخيريّة في الحد من مشكلة الفقر»، جامعة اليرموك، كلية الآداب، الأردن. 2014.

ثانياً: المراجع العربية المترجمة:

- *Saket, Gh. & Adas, D. (2015). Using the Decision Tree to Support the Decision to Grant Loans. Aleppo University Research Journal, Economic Sciences Series No. 14*

- *Abd Rabbo, M. (2013). Administrative Challenges Facing Charities and How to Address them: An Applied Study on a Sample of Charities in the Sultanate of Oman. Sultan Qaboos University Journal of Arts and Social Sciences, Sultanate of Oman.*
- *Miqdad, M. (2014). The Role of Palestinian Civil Associations in Reducing Poverty Rates in the Gaza Strip. The Islamic University, Gaza.*
- *Hilal, A. (2012). The Role of Charities in Reducing the Problem of Poverty. Yarmouk University, Faculty of Arts, Jordan. 2014.*

ثالثاً: المراجع باللغة الأجنبية:

- *Abbas, S. (2015). Deposit Subscribe Prediction using Data Mining Techniques based Real Marketing Dataset. International Journal of Computer Applications vol110*
- *Aksenova, S. (2004). Machine Learning with WEKA Explorer Tutorial for WEKA. Computer Science California State University, Sacramento California. P2*
- *Brunello, A. et al (2019). A Novel Decision Tree Approach for the Handling of Sequential and Time Series Data. Article J48SS, Computer MDPI*
- *Jože (2011). Fuzzy Decision Trees as Decision - Making Framework in the Public Sector. Yugoslav Journal of Operations Research, Number 2 205 - 224.*
- *Kabari, G. & Nwachukwu, E. (2013). Decision Support System Using Decision Tree and Neural Networks. Computer Engineering and Intelligent Systems, Vol. 4, No. 7*
- *Masud, R. & Rashedur, M. (2013). Decision Tree and naïve Bayes Algorithm for classification and Generation of Actionable Knowledge for Direct Marketing. Journal of software Engineering and Applications.*
- *Renukad et al (2012). Real Life Applications of Fuzzy Decision Tree. International Journal of Computer Applications (0975 - 8887).*
- *Vaidehi, R. (2016). Predictive Modeling to Improve Success Rate of Bank Direct Marketing Campaign. International Journal of Management & Business Studies, Vol6.*
- *Wittenn, I. et al (2011). Data Mining Practical Machine Learning Tools and Techniques. 3rdED, Morgan KauFmann, P5.*