

ABSTRACT

Data mining works on extraction of new knowledge from data, one of its main techniques is clustering. Which divide data in to groups. In this research, we want to know whether the classification of poverty that followed by the General Secretariat of the Zakat Chamber is optimal or there are other better classifications. K-means clustering method has been used to find clusters in Zakat Chamber poor people data set.

Zakat institute divide poor people into 3 groups. Zakat institute data set for the year 2011 has been used in this research. In all of the experiments with small number of clusters (3-7 clusters) the third group doesn't appear. The third group appear only when the number clusters set to 8 or more. These results of clustering are different from grouping adopted by the Zakat Chamber. This indicates that the Zakat Chamber classification system needs more analysis.

المستخلص

تنقيب البيانات يعمل علي إستخلاص معرفه جديدة من البيانات، إحدى تقنياته الرئيسية هي(العنقدة) والتي تعني بعملية تقسيم البيانات الي مجموعات. في هذا البحث نريد أن نعرف ما إذا كان تصنيف الفقر الذي تتبعه الأمانة العامة لديوان الزكاة مثالي ام أن هناك تصنيف آخر افضل منه. في هذا البحث طبقت إحدى طرق العنقدة وهي خوارزمية العنقدة K-means لايجاد تصنيف لفئات الفقراء السودانيين حسب بيانات ديوان الزكاة، استخدمت العنقدة للوصول لتصنيفات جديدة لبيانات الفقراء السودانيين. أفرز البحث عناقيد (مجموعات) علي تباين كبير مع التصنيفات المستخدمة بواسطة الديوان حاليا مما يستدعي إعادة النظر في تصنيف فئات الفقر الذي تتبعه الامانة العامة لديوان الزكاة. يقسم الديوان الفقراء لثلاثة مجموعات، لكن معظم النتائج تظهر المجموعة الاولى و الثانية (التجارب التي استهدفت انتاج 3-7 عناقيد).اما الفئة الثالثة (الاقل ظهورا) فلم تظهر إلا في التجارب التي استهدفت انتاج 8 عناقيد او اكثر. اعتمد البحث علي بيانات الحصر الشامل للفقراء في معهد علوم الزكاة- للعام 2011.

إهداء

اهدي هذا العمل المتواضع إلى أبي الذي لم يبخل علي يوماً بشيء

وإلى أمي التي نودتني بالحنان والمحبة

أقول لهم: أنتم وهبتموني الحياة والأمل والنشأة على شغف الاطلاع والمعرفة

وإلى زوجتي التي صبرت وعاشت مشواري والي وأسرتي جميعاً

ثم إلى كل من علمني حرفاً أصبح سنا برقه يضيء الطريق أمامي

1	1. الباب الاول المقدمة
2	1.1. مشكلة البحث
2	1.2. أهداف البحث
2	1.3. منهجية البحث
3	2. الباب الثاني مراجعة الادبيات
3	2.1 ماذا يعني التنقيب عن البيانات
4	2.2 مراحل اكتشاف المعرفة
4	2.3 تتضمن عملية اكتشاف المعرفة
5	2.4 التنقيب في البيانات
6	2.5 الأدوات الأساسية لتنقيب البيانات Basic Tools of Data Mining
7	2.6 اهداف للتنقيب في البيانات
7	2.6.1 هناك ثلاثة أهداف للتنقيب في البيانات
9	2.7 العنقدة
9	2.8 خوارزمية العنقدة K-Means Clustering
11	2.9 تحديد حزمة البرمجيات
11	WEKA 2.10
12	2.10.1 وصف التطبيق
15	2.10.2 العنقدة في تطبيق Weka
15	2.10.3 تنقية البيانات :
16	3. الباب الثالث: منهجية البحث والبيانات
16	3.1 المقدمة
16	3.2 اهداف الحصر
17	3.3 تصميم الاستمارة
18	3.4 الحصر الفعلي
18	3.5 المشاكل التي واجهت عمليات الحصر
18	3.6 العمليات الفنية ومعالجة البيانات
19	3.7 مرحلة بناء الاوزارن وحساب الدرجات للاسرة الفقيرة
20	3.8 كيفية إعداد الاوزان
26	3.9 تحليل البيانات
27	4. الباب الرابع المناقشة والنتائج

27	4.1 وصف البيانات
28	4.2 وصف التجارب
30	4.3 التجربة الاولى
31	4.3.1 تحليل النتائج
31	4.3.2 ملخص التجربة
32	4.4 التجربة الثانية
32	4.4.1 تحليل النتائج
32	4.4.2 ملخص التجربة
33	4.5 التجربة الثالثة
33	4.5.1 تحليل النتائج
34	4.5.2 ملخص التجربة
34	4.6 التجربة الرابعة
34	4.6.1 تحليل النتائج
35	4.6.2 ملخص التجربة
36	4.7 التجربة الخامسة
37	4.7.1 تحليل النتائج
37	4.7.2 ملخص التجربة
38	4.8 التجربة السادسة
39	4.8.1 تحليل النتائج
39	4.8.2 ملخص التجربة
40	5. الباب الخامس الخلاصة والتوصيات
40	5.1 الخلاصة
41	5.2 التوصيات

قائمة الاشكال

1. الشكل 2.1 استخراج نموذج البيانات والمهام.....8
2. الشكل 2.2 الصفحة الرئيسية لتطبيق weka النسخة 3.6.9.....12

3. الشكل 4.1 تصور لخصائص البيانات وتوزيعها.....29
4. الشكل 4.2 صورة لصفحة التجميع في برنامج الويكا.....30
5. الشكل 4.3 عناقيد التجربة.....31

- جدول رقم (3.1) يوضح تصنيف الفقراء حسب درجاتهم.....20
- جدول رقم (3.2) يوضح درجات عدد افراد الاسرة منسوبا الي عدد الغرف.....20
- جدول رقم (3.3) يوضح مادة بناء المنزل.....21
- جدول رقم (3.4) يوضح ملكية المسكن.....21
- جدول رقم (3.5) يوضح الصرف الصحي.....22
- جدول رقم (3.6) يوضح مياه الشرب.....23
- جدول رقم (3.7) يوضح نوع الاضاءة.....23
- جدول رقم (3.8) يوضح وقود الطهي.....24
- جدول رقم (3.9) يوضح التكيف مع الفقر.....24
- جدول رقم (3.10) يوضح بعد الصحة.....25
- جدول رقم (3.11) يوضح عدد افراد الاسرة.....25
- جدول رقم (3.12) يوضح التعليم لرب الاسرة.....25
- جدول رقم (3.13) يوضح قيمة الاستهلاك.....26
- جدول رقم (4.1) يوضح وصف البيانات المستخدمة.....27
- جدول رقم (4.2) يوضح نتيجة التجربة الاولى.....31
- جدول رقم (4.3) يوضح نتيجة التجربة الثانية.....32
- جدول رقم (4.4) يوضح نتيجة التجربة الثالثة.....33
- جدول رقم (4.5) يوضح نتيجة التجربة الرابعة.....34
- جدول رقم (4.6) يوضح نتيجة التجربة الخامسة.....36
- جدول رقم (4.7) يوضح نتيجة التجربة السادسة.....38