

وزارة التعليم العالي والبحث العلمي
الجامعة المستنصرية
كلية الادارة والاقتصاد

إستخدام التحليل العنقودي في عمليات التصنيف مع تطبيق عملي

إطروحة مقدمة الى

مجلس كلية الادارة والاقتصاد في الجامعة المستنصرية وهي جزء من متطلبات نيل درجة الدكتوراه

فلسفة في الاحصاء

مقدمة من قبل

وفاء سيد حسنين الشيخ

بإشراف الدكتور

آيدن حسن حسين الكناني

٢٠٠٦م

بغداد

١٤٢٧هـ

المستخلص

إن الهدف الأساس من التحليل العنقودي (العنقدة) هو تصنيف مجموعة من عناصر البيانات الى عناقيد ، بحيث تكون العناصر داخل العنقود أكثر تشابهاً مع بعضها البعض من العناصر في العناقيد الأخرى . ويمكن التعبير عن مفهوم التشابه بطرائق مختلفة وفقاً لهدف الدراسة والفرضيات الخاصة والمعلومات المسبقة للمشكلة .

إن أساليب العنقدة تقع ضمن فئتين رئيسيتين هما ، " العنقدة الحادة أو الواضحة " والتي تعني إنتماء العنصر الى عنقود واحد فقط بدرجة عضوية مساوية الى الواحد الصحيح وحدود واضحة بين العناقيد و" العنقدة الضبابية " والتي تعني إنتماء العنصر الى كل العناقيد مع درجات عضوية مختلفة محصورة بين الصفر والواحد .

يهدف البحث الى دراسة طرائق العنقدة الحادة ومن ثم دراسة طرائق العنقدة الضبابية وإقتراح طريقة في العنقدة الضبابية تعتمد على إيجاد قوة الارتباط بين عناصر البيانات ومراكز العناقيد المصنفة اليها من خلال إستخدام مقياس مسافة الارتباط ، بالإضافة الى دراسة طرائق صحة وقانونية العنقود .

وأخيراً تم إستخدام المحاكاة لتوليد بيانات تتبع التوزيع الطبيعي الخاصة بمصفوفة المتغيرات ، وبيانات تتبع التوزيع المنتظم القياسي الخاصة بمصفوفة عناصر التجزئة ، ومن ثم المقارنة بين طرائق العنقدة الحادة وتوصلت الدراسة الى أن طريقة (Ward) هي أفضل هذه الطرائق . كما تمت المقارنة بين طرائق العنقدة الضبابية وتوصلت الدراسة الى أن الطريقة المقترحة هي أفضل هذه الطرائق من خلال تصغير دالة الهدف ، بالإضافة الى ذلك تمت المقارنة بين الطرائق العنقودية الحادة والضبابية لإختيار العدد الصحيح من العناقيد من خلال إستخدام طرائق صحة وقانونية العنقود وتوصلت الدراسة الى أن الطريقة المقترحة هي أفضل هذه الطرائق من خلال تشخيصها العدد الأمثل من العناقيد .