توظيف تطبيقات التنقيب عن البيانات لتحليل درجات طلبة قسم المعلومات والمكتبات

م. علي الحر لازم / الجامعة المستنصرية/ كلية التربية الأساسية/ قسم الحاسبات
Dr.alialhur@uomustansiriah.edu.iq

المستخلص:
يهدف البحث الحالي الى التعرف على الأنماط السائدة في مستوى التحصيل الدراسي لطلاب قسم المعلومات والمكتبات للأعوام الدراسية (2006-2007) ولغاية العام الدراسي (2010-2011) وذلك لتعذر الحصول على درجات الطلبة الخريجين للأوقات التي تسبق هذا التاريخ وكذلك اكتمال المعلومات الخاصة بالطلبة ضمن العام الدراسي الأخير. وقد استخدم البحث المنهج التحليلي باستخدام تقنية التنقيب عن البيانات من خلال استخدام أسلوب العنقدة مستعينين ببرنامج ***weka*** للمساعدة في الحصول على النتائج النهائية.

المصطلحات الرئيسة / التنقيب عن البيانات، المعلومات والمكتبات.

1.1. مشكلة البحث وأهميته

يعد قسم المعلومات والمكتبات من الأقسام العريقة في كلية الآداب في الجامعة المستنصرية، اذ يعود تأسيسه الى اكثر من أربعة عقود من الزمن، وعلى الرغم من عراقة هذا القسم الا اننا نجد ان عملية الاقبال عليه قليله مقارنة مع اقسام أخرى مناظرة له في الكلية او في الكليات الأخرى، وقد لوحظ ذلك بشكل جلي في ضعف الاقبال عليه في الدراسات المسائية من قبل الطلبة، مقارنة بالأقسام الأخرى ضمن خطط القبول في الكلية، لذلك يتوجب التعرف على ميول الطلبة ونتائجهم في المقررات المدروسة في القسم بغية التعرف على مواطن الضعف في بعض المقررات مقارنة مع مقررات أخرى لنفس الطلبة، ولاجراء هذه العملية يتوجب ان تكون هناك قاعدة بيانات او مستودع بيانات رقمي يشمل العمليات التفصيلية التي تخص نتائج الطلبة في جميع المقررات الدراسية، للتعرف على مواطن الضعف والقوة لديهم، بغية تحسين الأداء في اختيار مفردات المواد ضمن المقرر الواحد او رفع بعض المقررات او إضافة أخرى، لرفع كفاءة القسم من قبل متخذي القرار، والذين بدورهم لن يتمكنوا من اجراء هذا التعديل دون الوقوف على نتائج محددة لسنوات متراكمة. ومن هنا تبرز مشكلة الدراسة الحالية في عدم وجود دراسة خاصة بنتائج الطلبة لاعوام طويلة، وكذلك ضعف قاعدة البيانات المتوفرة لدى القسم حول بيانات طلبتهم من حيث التنظيم وسهولة الوصول الى المعلومة المطلوبة ضمن الكم الكبير من البيانات المتراكمة لدى القسم حول درجات الطلبة وفرزها بالصورة الصحيحة بغية الوصول الى اهداف محددة.

2.1. فرضية البحث

في ظل المشكلة المطروحة يمكن صياغة الفرضية الآتية:
- توجد فروق بين التحصيل الدراسي للمقررات الخاصة بالمكتبات مقارنة بالمعلومات ولصالح المقررات الخاصة بالمعلومات.

1-3- أهداف البحث

1- دراسة مشكلة التراكم في بيانات الطلبة وإيجاد مداخل متعددة للوصول الى نتائج محددة.

2-بناء مستودع رقمي للبيانات الخاصة بدرجات الطلاب الاكاديمية للمراحل الدراسية الأربعة تكون مساندة لأصحاب القرار في الكلية للتعرف نتائج الطلبة الاكاديمية.

3- التعرف على العوامل المؤدية الى ضعف التحصيل الدراسي للطلبة في مقررات معينة دون غيرها باستخدام التنقيب عن البيانات.

1-4-حدود البحث

1-4-1- الحدود المكانية: قسم المعلومات والمكتبات-كلية الآداب –الجامعة المستنصرية.

1-4-2- الحدود الزمانية: الأعوام الدراسية 2006-2007 الى 2010-2011 (الدرجات الكاملة التي حصل عليها الباحث).

1-4-3- الحدود الموضوعية: سجلات درجات الطلبة في القسم المذكور للاعوام الدراسية السابقة.

1-5- منهجية البحث وأدواته

من خلال الأهداف المرسومة في البحث والتي يتوجب تحقيقها، تم استخدام المنهج المسحي للاطلاع على سجلات الطلبة، وكذلك استخدام المنهج التجريبي ممثلاً في استخدام برنامج الويكا للتنقيب عن البيانات باستخدام اسلوبي العنقدة وشجرة القرارات، مستعينين بالوثائق الخاصة بالدرجات وكذلك الملاحظة الشخصية للباحث كأدوات لجمع البيانات.

1-6- مجتمع الدراسة

نظراً لطبيعة البحث، فلم يتم اخذ مجتمع للدراسة او عينة وذلك لكون البحث لم يختص بالصفات البشرية للفئة المبحوثة وإنما اختص بدرجات أولئك الطلبة بصورة عامة وللفئات المتخرجة فقط للأعوام الدراسية السابقة الذكر.

1-7- الدراسات السابقة

1-7-1-دراسة زكريا مطلك الدوري، داليا عبد الحسين احمد، دور تنقيب البيانات في زيادة أداء المنظمة" دراسة تحليلية في المصرف الزراعي"، مجلة العلوم الاقتصادية والإدارية، مج13، ع48، 2007.

هدفت الدراسة الى تحسين أداء المصارف من خلال بناء نموذج يمكن ان يتنبأ بالمقترضين الموثوقين من غيرهم، اذ طبقت على المصرف الزراعي من خلال اخذ (277) قرضاً من أصل (1350) قرض لمدة خمس سنوات، وقد استخدم الباحثان المنهج التجريبي مستعينين بالمقابلات مع المسؤولين والملاحظة الشخصية من قبلهما كأدوات لجمع البيانات، أما عملية التنقيب فقد استخدم برنامج يعمل في بيئة SQL سيرفر. وقد خلصت الدراسة الى مجموعة من النتائج أهمها: كلما كان عدد الاعمدة في قواعد البيانات اكثر كلما كان فرص الوصول الى نتائج محددة اكبر، ان اعلى معدلات صافي الربح في المصرف كان للعام 2001 حيث ازدادت نسبة التسديد الى 60%.

1-7-2-دراسة صباح محمد موسى، وصادق حسين عبد الحسين، اثر تقنية تنقيب البيانات في تطوير عملية صنع القرارات دراسة استطلاعية في وزارة التعليم العالي والبحث العلمي العراقية، مجلة الإدارة والاقتصاد للسنة 38، ع130، 2015.

هدفت الدراسة الى معرفة اثر استخدام تقنيات المستودعات الرقمية في تطوير عملية اتخاذ القرارات، اذ شملت عينة الدراسة 69 منتسب من هم بمستوى مدير قسم وشعبة، وتم اخذهم بصورة قصدية، وقد تم استخدام المنهج التجريبي مستعينين بالمقابلة والملاحظة كأدوات لجمع البيانات وكذلك تم استخدام البرنامج الاحصائي SPSS لتحليل النتائج، وتوصلت الدراسة الى جملة من النتائج أهمها: حاجة المؤسسات اىل وسائل الوصول الى المعلومات السريعة كلما كان عدد البيانات كبير، وكذلك أهمية استخدام تكنولوجيا المعلومات في الوصول الى المعلومات المحددة.

1-8- موازنة البحث الحالي مع الدراسات السابقة:

من خلال الإطلاع على الدراستين السابقتين، فقد تشابه البحث الحالي معهما في جوانب واختلف معهما في جوانب أخرى، وهي، من خلال الهدف اذ هدفت دراسة زكريا وداليا الى تحسين أداء المصارف من خلال بناء نموذج يمكن ان يتنبأ بالمقترضين الموثوقين من غيرهم، ودراسة صباح وصادق الى معرفة اثر استخدام تقنيات المستودعات الرقمية في تطوير عملية اتخاذ القرارات، بينما كان هدف البحث الحالي هو دراسة مشكلة التراكم في بيانات الطلبة وإيجاد مداخل متعددة للوصول الى نتائج محددة من خلال بناء مستودع رقمي والتعرف على عوامل الضعف في بعض المقررات الدراسية دون سواها بالنسبة للطلبة في قسم المعلومات والمكتبات. وهو بذلك قد اختلف مع الدراستين السابقتين من حيث الهدف.

اما بالنسبة لمجتمع الدراسة فقد اشارت دراسة زكريا وداليا الى إتمام البحث حول المقترضين في المصارف ودراسة صباح وصادق الى متخذي القرار في وزارة التعليم العالي والبحث العلمي، وبذلك قد اختلف البحث الحالي معهما من حيث تركيزه على درجات الطلاب فقط دون التركيز على اشخاص محددين.

وقد اتبعت الدراستين السابقتين والبحث الحالي المنهج التجريبي إلا ان أدوات جمع البيانات كانت متشابهة بالنسبة من خلال استخدام المقابلة والملاحظة فيما اقتصر البحث الحالي على الملاحظة الشخصية فقط، من خلال الاطلاع المباشر على درجات الطلاب.

أما فيما يخص عمليات المعالجة فقد استخدمت دراسة صباح وداليا برنامج ضمن بيئة SQL سيرفر ودراسة صباح وصادق برنامج احصائي هو SPSS بينما ركز البحث الحالي على استخدام برنامج WEKA واستخدام اسلوبي العنقدة وشجرة القرارات.

1-2- التنقيب عن البيانات:

تمر عملية انتاج المعرفة بمراحل تبدأ أولها بوجود وتوافر بيانات معينة ومن ثم اجراء عملية معالجة عليها للتحول الى المعلومات، وكثرة توفر المعلومات بدورها تؤدي الى المعرفة والتي هي بأبسط تعريف " القدرة على تبرير معتقدات الشخص والمتأتية من كثرة المعلومات لديه "[1] ومن هنا تعد البيانات الصيغة الأولى للأشكال والصور والاصوات وغيرها.. قبل ان تتم عليها عملية المعالجة

[1]) علاء اطلس فواز. استراتيجيات التدريس الفعال . ط1. (بغداد: دار الموج للنشر، 2014). ص 25.

للتحول بدورها الى معلومات. ونظراً للتوسع الكبير في عالم الاتصالات الذي هو بدوره يعد السمة الأولى للعصر الحالي، فقد تدفقت وسائل المعلومات بكافة الاشكال والاوعية واللغات، مما جعل امام مؤسسات المعلومات ومنها المؤسسات الاكاديمية التعليمية، مشكلة السيطرة على هذه البيانات لغرض إجراء المعالجة الدقيقة لها للوصول الى المعلومات المطلوبة بكل دقة، اذ لم تعد النظم التقليدية والمتمثلة بقواعد البيانات كافية لاستخلاص المعلومات بدقة تناسب احتياجات المستفيدين منها، خاصة وبعد ان بدأ عصر مستودعات البيانات .Data Warehouse لذلك ظهرت الحاجة الملحة لاستخلاص المعلومات من بين الكم الهائل من البيانات، وبمعنى آخر هي عملية معالجة للبيانات واستخلاص المعلومات المطلوبة بنفس الوقت، وهي بالمصطلح الدقيق عملية التنقيب عن البيانات Data mining.والذي يمكن تعريفه على انه: عملية غربلة وتحليل كمية كبيرة من البيانات لاستخراج ما هو ذو معنى واكتشاف معرفة جديدة تهدف الى وصف الاتجاهات الماضية والتنبؤ بالمستقبل[1]. كما عرفت على انها: عمليات استخلاص المعلومات المتوقعة من قواعد البيانات الكبيرة الحجم مع الاكتشاف الاوتوماتيكي للمعرفة المخفية ضمن قواعد البيانات[2]. وهناك العديد من الكتاب الذي عرفوا عملية التنقيب عن البيانات الا انها بالمجمل تصب في معنى واحد ألا وهو استخلاص معلومات محددة من بين كم هائل من المعلومات للتنبؤ بالمستقبل.

2.1.1. فوائد عملية التنقيب عن البيانات:

تعد عملية التنقيب عن البيانات ومن خلال التعاريف السابقة ضرورية لالتمام المهام المتعلقة بإيجاد واستخلاص معلومات محددة من كميات كبيرة من البيانات بصورة عامة، إلا أن ابرز الفوائد لتلك العملية هي:

1-تمكن عملية التنقيب عن البيانات المستفيدين منها من الوصول الى معلومات لم يكن بالإمكان الوصول اليها بالطرق الأخرى.

2-تمكن أصحاب القرار من إيجاد نمط استنتاجي عن طريق فهم الماضي للوصول الى تنبؤ حول مستقبل المسألة المطلوبة.

3- تمكن عملية التنقيب عن البيانات المستفيدين منها معرفة واكتشاف الارتباطات والاتجاهات والانماط السائدة في عمل مؤسسة معينة[3].

2.1.2. مراحل التنقيب عن البيانات:

تمر عملية التنقيب عن البيانات بمراحل، هي:

1-**فهم طبيعة عمل المؤسسة**: تتلخص المرحلة الأولى بعملية فهم عمل المؤسسة الام والتي هي تحدد نوع المطلوبة.

2-فهم البيانات: وغي هذه المرحلة تتم عملية معرفة نوع البيانات المخزنة ومسارها.

---

[1]) اواب الجناعي، وآخرون. استكشاف بعض الأنماط المؤثرة في الأداء الاكاديمي باستخدام تقنيات التنقيب عن البيانات. مجلة العلوم والتكنولوجيا. المجلد 16، العدد 1، 2011. ص 25.

[2]) صادق عبد العزيز مهدي. نظم إدارة قواعد البيانات. (بغداد: دار الفتح، 2016). ص33.

[3]) زكريا مطلك الدوري، داليا عبد الحسين احمد. دور التنقيب عن البيانات في زيادة أداء المنظمات . مجلة العلوم الاقتصادية. المجلد 12، العدد 48، 2007. ص 47.

3-اعداد البيانات: وفي هذه المرحلة تتم عملية تنظيم البيانات وضمان تكاملها وخلوها من مشكلة التكرار في البيانات وخلو بعض القيود من معلومات.

4-التنقيب عن البيانات: وفي هذه المرحلة تتم عملية تحليل البيانات للوصول الى النتيجة المطلوبة[1].

5- التقييم: وغي هذه المرحلة تتم عملية تقييم النتائج التي توصلت اليها عملية التنقيب.

2-1-3-طرق التنقيب عن البيانات:

هناك العديد من طرق التنقيب عن البيانات، هي:

1-تحليل الارتباط.

2-الخوارزميات الجينية.

3-شبكات النظرية الافتراضية.

4-مسار المجموعة الخام.

5-الشبكات العصبية.

6-التحليل الاحصائي.

7-شجرة القرارات.

2-2- تنقيب البيانات في المؤسسات التعليمية:

شهد عام 1989 ظهور اول مصطلح خاص بعملية التنقيب عن البيانات، وذلك من خلال ورشة عمل عقدت في لندن، مما اثار هذا الموضوع انتباه الباحثين والمتخصصين في مجال المعلومات، مما أدى الى انعقاد مؤتمر عام 1994 واصبح يعقد بصفة دورية كل عام، ولم تكن مؤسسات المعلومات والتعليمية منها بعيدة عن هذا المجال اذ كان العام 1998 بداية استخدام هذا الجانب في تحليل البيانات الخاصة بالمستفيدين منها، وغيرها من العمليات ذات الصلة بتحسين اداءها والتنبؤ بالاحتياجات المستقبلية التي ستضطلع بها[2].

اذ ان الانتشار الواسع في أنظمة التعليم الجديدة كالتعلم الالكتروني والتعليم عن بعد، أدى الى ظهور التنقيب عن البيانات التعليمية (educational data mining) وتحليل التعلم (learning analysis)، وكلا المفهومان يهتمان بعمليات جمع وتحليل وتفسير النتائج المستخرجة، ومن الأسباب الموجبة لاستخدام هاتين التقنيتين، هي:

---

[1]) صباح موسى محمد . اثر تقنية التنقيب عن البيانات في تطوير عملية صنع القرارات . دراسة استطلاعية في وزارة التعليم العالي والبحث العلمي العراقية. مجلة الإدارة والاقتصاد. العدد 103، 2015. ص93.

[2]) محمد اطلس فيروز. استخدام تقنية التنقيب عن البيانات لدعم الجودة الشاملة باستخدام SPSS في معهد إدارة الاعمال. أطروحة دكتوراه . سوريا. المعهد العالي للعلوم التطبيقية والتكنولوجيا. 2017.

1-انتشار استخدام نظم المعلومات في التعليم، اذ إننا نشهد في الوقت الحالي تعزيز لهذا الدور من خلال حث التدريسيين على نشر المحاضرات والدرجات الخاصة بالطلاب ورفع المحاضرات والتبليغات على البروفايلات الاكاديمية للتدريسيين.

2-سهولة توحيد الصيغ المستخدمة في نشر المعلومات ، اذ اصبح من السهل تحويل النتائج الخاصة بالطلبة عن طريق ملفات الاكسل او تحويلها من صيغ أخرى الى الصيغة المعنية.

3-تقنيات المعالجة من خلال برامج الحوسبة، اذ اصبح الوقت الحالي يشهد انتشار العديد من برامج الحوسبة الخاصة بالمعالجات الإحصائية للنتائج مثل برامج SPSS وغيرها.

4- تطور التقنيات المستخدمة في التنقيب عن البيانات، وكما مر ذكرها في الصفحات السابقة.

3.1.الجانب العملي:

تم في هذا الجزء اجراء العمليات الخاصة بتحليل النتائج الخاصة بالطلبة تبعا لسنوات البحث المذكورة سابقاً، ومن أجل الإيفاء بهذه المتطلبات تم تقسيم الجانب العملي الى عدة مراحل، هي:

1-مرحلة جمع البيانات ومعالجتها: وقد تم في هذه المرحلة اخذ البيانات الخاصة بنتائج الطلاب وتقسيمها الى مواد تخصصية بالمكتبات وأخرى بالمعلومات، اذ شملت المواد الخاصة بعلم المعلومات (6) وهي ((الحاسوب، في المرحلة الأولى)، (نظم معلومات وقواعد بيانات في المرحلة الثانية)، (علم المعلومات، الحوسبة في المرحلة الثالثة)، (شبكات، خزن واسترجاع المعلومات في المرحلة الرابعة))، بينما المواد الخاصة بعلم المكتبات (16) وهي (تنظيم اوعية معلومات وصفي وموضوعي، مراجع عربية، المدخل الى علم المكتبات في المرحلة الأولى)، (تنظيم اوعية معلومات وصفي وموضوعي ج2، إدارة مؤسسات، مراجع ، وثائق، للمرحلة الثانية)،(مصادر معلومات، نصوص مكتبية، نظم تصنيف، للمرحلة الثالثة)، (مصادر، بناء مجموعات، ببلوغرافية، مخطوطات للمرحلة الرابعة)، إضافة الى عدد من المواد تعتبر ساندة مثل حقوق انسان وديمقراطية، واللغتين العربية والإنكليزية، كذلك تم دمج درجات المواد المتشابهة مثل تنظيم اوعية معلومات للمرحلتين الأولى والثانية باعتبارها نفس المادة وغيرها من المواد المتكررة. إذ لم تخلو هذه المرحلة من بعض المشكلات اذ ظهر من خلال عمليات فرز البيانات هناك تشابه بالاسماء الثلاثية لبعض الطلبة وخصوصا ان البحث قد اعتمد على الاسم كمدخل أساس لعدم توفر ارقام إحصائية للطلبة مثبتة ضمن البيانات الموجودة وكذلك حداثة استخدام الأرقام الإحصائية بالنسبة للبروفايلات الاكاديمية للطبة. وكذلك مشكلة بعض الطلاب العائدين للعام الدراسي بمواد معينة فقط مما اثر في عملية تجميع البيانات كونه قد درس مواد واجتازها في عام دراسي معين ومواد أخرى في عام دراسي آخر, وكذلك ظهرت مشكلة أخرى في كتابة بعض الدرجات الخاصة بالطلاب اذ تم اعتماد الفارزة لكتابة الأجزاء العشرية المكملة للدرجة واحيان أخرى تم كتابة الفاصلة على شكل نقطة، كما تم توصيف الأرقام الى قيم كتابية مثل (ضعيف، مقبول، متوسط، جيد، جيد جداً، امتياز) لتكون مرادفة لدرجات الطلاب في جميع المواد.

ت	المادة	البدائل الكتابية					
		ضعيف	مقبول	متوسط	جيد	جيد جدا	امتياز
1		49-1	59-50	69-60	79-70	89-80	100-90
2		49-1	59-50	69-60	79-70	89-80	100-90
3		49-1	59-50	69-60	79-70	89-80	100-90
4		49-1	59-50	69-60	79-70	89-80	100-90
5		49-1	59-50	69-60	79-70	89-80	100-90
6		49-1	59-50	69-60	79-70	89-80	100-90

2-مرحلة تحويل صيغة الملفات الى صيغة مقبولة لدى برنامج WEKA:

أي انه تم تحويل من صيغة excel الى صيغة Arrf، اذ شملت هذه العملية تحويل البيانات فقد كانت مخزنة لدى القسم بصيغة الاكسل، وهي الصيغة المعتادة في عمليات الحفظ لدى الأقسام، فتوجب عملية تحويلها لصيغة مفهومة لدى البرنامج، اذ بدأت أولى الخطوات بفتح الملف الأساسي ومن ثم حفظه بصيغة CSV وفي مكان تخزيني اخر في الجهاز، ومن ثم فتح البرنامج WEKA وبدأ عملية استيراد الملفات من خلال القائمة Tools وبامتداد ARRf لكي يصبح ضمن القواعد المحلية للبرنامج.

3- توصيف بيانات الطلبة:

اذ شملت هذه المرحلة تقسيم البيانات في شكلها الاولي الى صفوف مثلت المواد الدراسية للمقررات الدراسية واعمدة مثلت العوام الدراسية(الفترة الزمنية الخاصة بالبحث).

4-المصفوفات الخوارزمية المستخدمة بالبرنامج :

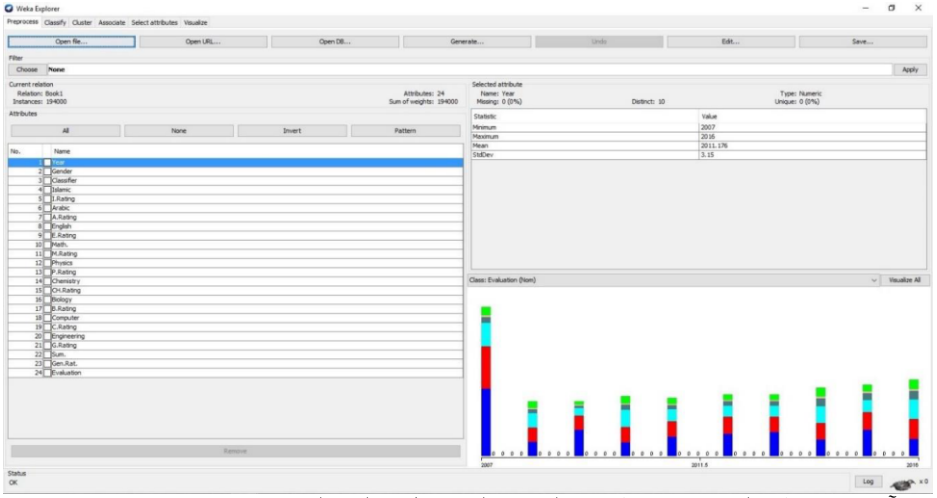
مثل ما هو معلوم توجد العديد من الخوارزميات داخل برنامج الويكا الا ان في بحثنا هذا استخدمت خوارزميات شجرة القرارات والعنقدة.

5-**ادخال المعلومات وتحليلها من قبل البرنامج،** والتي مرت بالخطوات الآتية:

5-1-تشغيل البرنامج من اجل عرض الخيارات المتوافرة داخل البرنامج.

5-2-وبعد ان يتم استيراد البيانات على شكل الملف المطلوب، والذي قمنا بتحويله بالخطوات السابقة بعد تحويله من شكلة الأصلي (الاكسل) وباخذ الخيار المستكشف.

5-3- تنفيذ الامر الخاص بالخوارزمية K-means ، وهنا يقوم البرنامج بتجميع المعلومات.



والآن نضغط على خيار اظهار النتائج لنحصل على النتيجة:

Missing values globally replaced with mean/mode

Cluster centroids: Cluster#

Attribute Full Data 0 1 2 3 4

(194000) (32549) (37750) (26548) (50159) (46994)

Year 2011.1762 2010.9052 2019.9026 2018.6187 2007.4365 2006.9284

Catalog S 70.2545 78.5594 75.2042 26.9785 77.4828 77.2587

CatalgD 54.9226 64.7201 55.3481 24.23 61.0337 58.6111

Storge 64.1653 69.306 71.227 29.7461 68.6999 69.5365

C0lication 70.2545 78.5594 75.2042 26.9785 77.4828 77.2587

Information . 38.1649 43.6222 40.5999 14.2985 41.5857 42.2606

database 36.8234 40.9457 38.9215 21.0148 37.8142 40.1558

mar 70.2545 78.5594 75.2042 26.9785 77.4828 77.2587

enter 55.6278 59.6339 61.1499 23.8461 60.5171 61.1531

Bibilio 31.0314 0.0053 67.557 14.885 0.0004 65.4223

Computer 13.797 28.9897 0 6.2846 31.2246 0

Manegment 16.3933 35.312 0 10.0384 35.1768 0

computer 54.9226 64.7201 55.3481 24.23 61.0337 58.6111

connection 70.2545 78.5594 75.2042 26.9785 77.4828 77.2587

Sum. 381.1803 421.0944 410.0075 171.322 413.5352 414.3981

Time taken to build model (full training data) : 135.12 seconds

= Model and evaluation on training set =#=

Clustered Instances

ت	المادة	النسبة المؤية	صنف المادة
1	تنظيم اوعية معلومات وصفي 1و2	*3.11%*	مكتبات
2	تنظيم اوعية معلومات موضوعي 1و2	*2.78%*	مكتبات
3	مراجع عربية وانكليزية	*3.01%*	مكتبات
4	حاسوب عام	*4.71%*	معلومات
5	إدارة مؤسسات	*8.33%*	مكتبات
6	وثائق	*6.66%*	مكتبات
7	نظم معلومات وقواعد بيانات	*9.11%*	معلومات
8	اتصالات	*4.09%*	معلومات
9	علم المعلومات	*7.61%*	معلومات
10	مصادر معلومات1 و2	*5.44%*	مكتبات

11	نصوص مكتبية	*5.55%*	مكتبات
12	حوسبة	*7.62%*	معلومات
13	خزن واسترجاع المعلومات	*6.12%*	معلومات
14	ببلوغرافيات	*9.42%*	مكتبات
15	شبكات	*9.18%*	مكتبات
16	مخطوطات	*7.16%*	مكتبات

من ملاحظة الجدول أعلاه يتبين ان المواد المحللة كانت بالمجموع (16) مادة منها (10) مواد خاصة بالمكتبات و (6) خاصة بالمعلومات.

بلغت مجموع الدرجات المستحصلة لمواد المكتبات (61.12%) من مجموع الدرجات وبمعدل بلغ (6.11) للمادة الواحدة ضمن (10) مواد الخاصة بعلم المكتبات من مجموع (16) مادة الكلية، بينما كانت نسبة الدرجات المستحصلة لمواد المعلومات (38.69%) وبمعدل (6.44) للمادة الواحدة ضمن (6) مواد الخاصة بعلم المعلومات ضمن (16) مادة الكلية التي تم تحليلها.

النتائج:

من خلال استخراج النتائج السابقة، يمكن تلخيص اهم النقاط الأساسية للنتائج:

1-اظهر البحث ان هناك تكرار في المواد الدراسية ضمن المقررات الدراسية لاقسام المعلومات والمكتبات لأكثر من مرحلة دراسية واحدة او مقرر واحد، مثل تنظيم اوعية وصفية 1و2 وموضوعية 1 و 2 وغيرها.

2-اشارت النتائج الى ان البيانات الخاصة بالطلبة لم تكن مرتبة ومنظمة بطريقة منهجية لدى القائمين على الأقسام سوى بيانات مخزنة لاستخراج النتائج السنوية للطلبة وكما يطلب منهم في اقسام التسجيل.

3-تبين ومن خلال النتائج ان المواد المدرسة لاقسام المكتبات في المراحل الدراسية الأربعة قد ركزت على علم المكتبات من خلال (23) مادة مقابل (6) فقط خاصة بعلم المعلومات.

4- ظهور فوارق طفيفة بين نتائج الطلبة في درجاتهم بين علم المعلومات والمكتبات ولصالح علم المعلومات بنسبة وصلت الى (6.44) مقابل (6.11) لمواد علم المكتبات.

5- إمكانية تطبيق برنامج الويكا في تحليل النتائج وسهولة استخدامه وتنصيبه وسهولة تحويل البيانات المخزنة بصيغة اكسل الى صيغ أخرى يمكن للبرنامج قراءتها ومعالجتها.

التوصيات:

من خلال النتائج السابقة يوصي الباحث بالآتي:

1-تطبيق برنامج الويكا ولخوارزميات أخرى في تحليل البيانات الخاصة بمؤسسات المعلومات لاستخلاص نتائج سريعة.

2-التركيز على طرق وأساليب التنقيب عن البيانات كونها تستطيع تحليل نتائج كبيرة جداً في أوقات قياسية، باستخدام النظم الإحصائية مثل SPSS او من خلال استخدام البرمجة المرئية مثل استخدام برنامجي SQL Server او Visual Basic.

3-القيام بعمل موازنة بين المواد والمقررات الدراسية المدرسة للطلبة بين علمي المكتبات والمعلومات.

4- تحديث المناهج والمفردات الدراسية بإضافة مواضيع جديدة للمقررات مثل علم البيانات، وتحليل النظم، وغيرها من المفردات التي تضيف معلومات جديدة لدى الطلبة.

المصادر:

1. Awab Al-Jana'i, et al. Explore some influential patterns in academic performance using data mining techniques. Journal of Science and Technology. Volume 16, Number 1, 2011. p. 25.
2. Zakaria Mutlak al-Douri, Dalia Abdul Hussein Ahmed. The role of data mining in increasing the performance of organizations. Journal of Economic Sciences. Volume 12, Number 48, 2007. p. 47.
3. Sadiq Abdul Aziz Mahdi. Database management systems. (Baghdad: Dar al-Fath, 2016). P. 33.
4-Sabah Musa Mohammed. The impact of data mining technology in the development of decision-making. An exploratory study in the Iraqi Ministry of Higher Education and Scientific Research. Journal of Management and Economics. Number 103, 2015. p. 93.
5. Alaa Atlas Fawaz. Effective teaching strategies. I 1. (Baghdad: The Wave House, 2014). P. 25.
6. Mohamed Atlas Fayrouz. Using data mining technology to support total quality using SPSS in the Institute of Business Administration. PhD thesis. Syria. Higher Institute of Applied Sciences and Technology. 2017.

توظيف تطبيقات التنقيب عن البيانات لتحليل درجات طلبة قسم المعلومات والمكتبات

م. علي الحر لازم

Recruitment Data mining Application to Analysis the information and library department students degrees

Abstract:

This research aims at identifying the common patterns at the level of academic achievement for students of the Department of Information and Libraries for the academic years (2006-2007) to (2010-2011), due to the difficulty of getting the graduate students' grades before this date, as well as the accumulation of the students' information in 5he last academic year. 5 The research used the analytical method using data mining technique using clustering method using the weka program to help obtain the final results.

Keyword / Data maining , information and library department.