تطبيق آليات التنقيب عن البيانات في تحسين مستوى خدمة الإعارة:
دراسة تطبيقية على بيانات الإعارة للكبار بمكتبة مصر العامة

د. كريمان بكنام صدقي عبد العزيز
مدرس علم المعلومات
بقسم المكتبات والوثائق وتقنية المعلومات
كلية الآداب، جامعة القاهرة
kareman14@cu.edu.eg

تاريخ الاستلام: 23 سبتمبر 2024 **تاريخ القبول: 2 نوفمبر 2024**

المستخلص:

في ظل التطورات التكنولوجية التي يشهدها عصر تكنولوجيا المعلومات وتضخم مصادر المعلومات وتنوعها، وما تتيحه من بيانات ضخمة يصعب التعامل معها؛ ظهرت الحاجة إلى تطوير أساليب أكثر قدرة على إجراء عمليات البحث والتنقيب عن البيانات كتقنية تهدف إلى استنتاج المعرفة من كميات هائلة من البيانات، وذلك بتحويلها من بيانات متراكمة وغير مفهومة إلى معلومات قيمة يمكن استغلالها والاستفادة منها لبناء إستراتيجيات اتخاذ القرارات.

ومن هذا المنطلق، جاءت هذه الدراسة لتهدف إلى تطبيق أساليب التنقيب عن البيانات في تحليل بيانات ملفات الإعارة للكبار في مكتبة مصر العامة، وتقدم الدراسة بالاعتماد على المنهج الوصفي باستخدام الأسلوب التحليلي عرضا للمفاهيم الرئيسية حول التنقيب عن البيانات وتطبيقاته وقضاياه الأساسية ومهامه، وما الأساليب والطرق المستخدمة في التنقيب عن البيانات، وتوضيح ديناميكية عمل كل منها، إضافًة إلى استخدام التقنيات مفتوحة المصدر ومنها: برنامج Orange3 لاستخراج المعرفة من بيانات ملفات الإعارة للكبار في مكتبة مصر العامة، وكيفية الاستفادة من تلك المعرفة في الكشف عن احتياجاتهم البحثية وميولهم القرائية، وتقديم المقترحات حول تزويدهم بمصادر المعلومات التي تلبي تلك الاحتياجات، واستنباط رؤى جديدة لتنمية مجموعات المكتبة.

ومِن أبرز ما خَلُصت إليه تلك الدراسة أنَّه يجب توجيه المكتبات ومؤسسات المعلومات نحو الإفادة من تقنيات التنقيب عن البيانات في تقديم خدماتها بكفاءة وجودة عالية؛ وذلك لما توفره من معلومات حول المستفيدين منها وتحديد الاحتياجات القرائية لهم، مَما يُسهم في تقديم رؤية جديدة حول ما تقتنيه من مصادر معلومات وما تُقدمه من خدمات تلبي احتياجات تلك المستفيدين، كما يمكن الإفادة مِن أدوات التنقيب عن البيانات في بناء قاعدة بيانات تُساعد العاملين في المكتبات ومؤسسات المعلومات ومتخذي القرار من اتخاذ القرار السليم تجاه تنمية مجموعات المكتبة والخدمات المستقبلية التي يجب أن تقدمها.

الكلمات المفتاحية: التنقيب عن البياناتData Mining؛ الذكاء الاصطناعي؛ مكتبة مصر العامة؛ الاحتياجات القرائية والبحثية؛ برنامج Orange3.

أولا: الإطار المنهجي:

1/1 التمهيد:

تمرُ عملية إنتاج المعرفة بمراحل؛ تبدأ أولها بوجود وتوافر بيانات معينة في صورتها الأولية لم تؤدِّ إلى أي معنى؛ ومن ثَم إجراء عمليات المعالجة عليها لتتحول إلى معلومات ذات معنى ودلالات مفهومة، وكثرة توافر المعلومات بدورها تؤدي إلى إنتاج المعرفة؛ ومع ظهور كميات كبيرة من البيانات الضخمة التي تنمو وتزداد بشكل سريع قد يصعب أو يستحيل معالجتها باستخدام الطرق التقليدية؛ ظهرت الحاجة إلى تطوير أساليب أكثر قدرة على إجراء عمليات البحث والتنقيب عن البيانات كتقنية تهدف إلى استنتاج المعلومات والمعارف من تلك الكميات الهائلة من البيانات.

ويُعد التنقيب عن البياناتData Mining مجالًا فرعيًا متعدد التخصصات لعلوم الذكاء الاصطناعي Artificial Intelligence وعلوم الإحصاء والرياضيات Statistics، يهدف إلى استخراج المعلومات من مجموعات البيانات، وذلك بتحويلها من بيانات متراكمة وغير مفهومة إلى معلومات قيمة يمكن الاستفادة منها في بناء التنبؤات المستقبلية، وتحليل الاتجاهات القرائية والبحثية لدى المستفيدين؛ مما يسمح لمتخذي القرار اتخاذ القرارات الصحيحة في الوقت المناسب.

ونظرًا لما يُقدمه التنقيب عن البيانات من تقنيات تعمل على استنتاج المعرفة من كميات البيانات الضخمة؛ سعت العديد من المؤسسات إلى دمج تلك التقنيات في العمليات التي تقوم بها والخدمات التي تقدمها للمستفيدين منها، وكانت المكتبات ومؤسسات المعلومات كغيرها من المؤسسات التي استفادت من تطبيق تقنيات التنقيب عن البيانات في العديد من الخدمات؛ كالإعارة، والتزويد، والتصنيف، والخدمة المرجعية، والبحث عن المعلومات واسترجاعها؛ وذلك لتمكين مديري المكتبات والمؤسسات المعلوماتية من إعادة توزيع مصادر المعلومات والخدمات المقدمة بطريقة مختلفة وفقا لاحتياجات واهتمامات وميول المستفيدين منها (درويش،2012).

ومن هذا المنطلق تعمل الدراسة على كيفية تطبيق آليات التنقيب عن البيانات في تحليل بيانات ملفات الإعارة للكبار في مكتبة مصر العامة باستخدام برنامج Orange3 ، لاستخراج المعرفة من ملفات بيانات الإعارة للكبار في مكتبة مصر العامة، وتوضيح كيفية الاستفادة من تلك المعرفة في الكشف عن الاحتياجات القرائية والبحثية لهم؛ وذلك من خلال عرض للمفاهيم الرئيسية حول التنقيب عن البيانات وتطبيقاته وقضاياه الأساسية ومهامه، وما الأساليب والطرق المستخدمة في التنقيب عن البيانات، وتوضيح ديناميكية عمل كل منها، إضافًة إلى استخدام التقنيات مفتوحة المصدر، ومنها: برنامج Orange3 وبيان بنيته الهيكلية وآليات تطبيقه وتطبيقه في استخراج المعرفة من ملفات بيانات الإعارة للكبار في مكتبة مصر العامة، وكيفية الاستفادة من تلك المعرفة في تقديم المقترحات حول تزويدهم بمصادر المعلومات التي تلبي تلك الاحتياجات.

2/1 إشكالية الدراسة ومبرِّراتُها:

يُعد الذكاء الاصطناعيArtificial Intelligence وتطبيقاته واحدًا من العلوم التقنية الحديثة التي تهدف إلى قيام الآلة بكل عمل يقوم به الإنسان بطريقة أفضل وأكثر تيسيرًا؛ ونظرًا لما يُقدمه الذكاء الاصطناعي من تقنيات تدخل في العديد من المجالات ومنها مجال التنقيب عن البيانات، سعت العديد من المؤسسات كالمكتبات ومؤسسات المعلومات إلى تطبيق آليات التنقيب عن البيانات للمساعدة في تطوير أدوات عمل مبتكرة في ظل هذا الكم الهائل من تضخم مصادر المعلومات، لتُمكن أخصائي المكتبات والمعلومات من تقدير احتياجات المستفيدين تجاه مجموعاتها وإعادة النظر في اقتناء مصادر المعلومات والخدمات المقدمة لهم بما يتناسب مع تلبية احتياجاتهم البحثية وميولهم القرائية.

وتُعد مكتبة مصر العامة من أكبر المكتبات العامة في مصر والتي تخدم فئات متعددة من المستفيدين بأعمار مختلفة، للتشجيع على تنمية عادة القراءة والاستفادة من المواد الثقافية والتعليمية المتاحة بالمكتبة؛ ومع تضخم مصادر المعلومات وتنوعها، ظهرت الحاجة لتطوير أساليب أكثر قدرة على إجراء عمليات البحث والتنقيب عن البيانات في التعرف على الميول القرائية والاتجاهات البحثية للمستفيدين منها، وإتاحة المعلومات بالشكل الذي يُعزز من استخدامها في اكتشاف رؤى جديدة لمجموعات المكتبة.

وهنا تبرُز إشكالية الدراسة في تطبيق آليات التنقيب عن البيانات في تحليل بيانات ملفات الإعارة للكبار في مكتبة مصر العامة باستخدام برنامج Orange3 لاستخراج المعلومات من تلك الملفات، وتوضيح كيفية الاستفادة منها في الكشف عن الاحتياجات القرائية والبحثية للمستفيدين من المكتبة؛ وذلك لتقديم المقترحات حول تزويدهم بمصادر المعلومات التي تلبي تلك الاحتياجات ولتحسين كفاءة وجودة الخدمات المقدمة لهم.

3/1 أهداف الدراسة:

يتمثل الهدف الرئيسي للدراسة في" تطبيق آليات التنقيب عن البيانات في تحسين مستوى خدمة الإعارة للكبار في مكتبة مصر العامة" كتقنية تهدف إلى استنتاج المعرفة من الكميات الهائلة من البيانات، وإتاحة المعلومات بالشكل الذي يعزز من استخدامها في اكتشاف رؤى جديدة لمجموعات المكتبة، وتقديم الخدمات المعلوماتية التي تتعلق بتلبية تلك الاحتياجات؛ وينطوي هذا الهدف الرئيسي على مجموعة من الأهداف الفرعية سعت الدراسة إلى تحقيقها، وهى:

1- رصد التطبيق العملي لبرنامج Orange3؛ وبيان بنيته الهيكلية.
2- توضيح آليات تطبيق برنامج Orange3 في استخراج المعرفة من ملفات بيانات الإعارة للكبار في مكتبة مصر العامة.
3- عرض نتائج التنقيب عن بيانات الإعارة للكبار في مكتبة مصر العامة على متخذي القرار بالمكتبة؛ لتقديم المقترحات حول استنباط رؤى جديدة لما تقتنيه من مصادر معلومات، وما تُقدمه من خدمات تلبي احتياجات المستفيدين منها.

4/1 تساؤلات الدراسة:

لتحقيق أهداف الدراسة، تحاول الدراسة الإجابة عن التساؤلات التالية:

1- ما البنية الهيكلية لتطبيق Orange3؟
2- ما مراحل التطبيق العملي لبرنامج Orange3 في استخراج المعرفة من ملفات بيانات الإعارة للكبار في مكتبة مصر العامة؟
3- كيف سيساعد تطبيق تقنية التنقيب عن البيانات متخذي القرار في مكتبة مصر العامة لاتخاذ القرارات تجاه تنمية مجموعات المكتبة، والخدمات المستقبلية التي يمكن تقديمها للمستفيدين منها؟

5/1 حدود الدراسة:

- **الحدود الموضوعية:** تتناول الدراسة موضوع تطبيق آليات التنقيب عن البيانات في تحسين مستوى خدمة الإعارة للكبار باستخدام برنامج Orange3 المفتوح المصدر كأحد أدوات التنقيب عن البيانات، وبيان بنيته الهيكلية وآليات تطبيقه في استخراج المعرفة من ملفات بيانات الإعارة للكبار في مكتبة مصر العامة " الفرع الرئيسي بالدقي".
- **الحدود الزمنية:** تناولت الدراسة تحليل ملفات بيانات الإعارة من قسم الكبار في مكتبة مصر العامة؛ وذلك في الفترة من عام 2011- وحتى ديسمبر 2023.

6/1 منهج الدراسة وأدواتها:

فرضت طبيعة الدراسة – وما تَسعى إليه من أهداف– استخدام المنهج الوصفي باستخدام الأسلوب التحليلي في دراسة تطبيق آليات التنقيب عن البيانات في تحسين مستوى خدمة الإعارة للكبار في مكتبة مصر العامة؛ وذلك للمساعدة في تطوير أدوات عمل مبتكرة تُمكن أخصائي المكتبات والمعلومات من تقدير احتياجات المستفيدين تجاه المجموعات والخدمات المقدمة لتلبية احتياجاتهم القرائية والبحثية، كما تم الاعتماد على أداة تحليل المحتوى في تجميع المادة العلمية عن طريق الاطلاع على المصادر البحثية المُتاحة في الإنتاج الفكري المنشور باللغتين العربية والإنجليزية والأدلة الإرشادية لبرنامجOrange3 والمُتاحة على موقع النظام على شبكة الإنترنت، إضافًة إلى الفيديوهات التعليمية المُتاحة على اليوتيوب.

وقد اقتصرت الدراسة على تجميع ملفات الإعارة للكبار من مكتبة مصر العامة بفرعها الرئيسي بالدقي؛ وذلك لقبولهم التعاون والحصول على تلك الملفات بموافقة من مدير المكتبة ونائبه، والقائمين على سجلات الإعارة للاستفادة من تطبيق نتائج الدراسة؛ والقيام بمجموعة من الإجراءات البحثيَّة لإتمام الدراسة؛ وهي كالآتي:

1- الاطلاع على العديدِ من الإنتاج الفكري العربي والأجنبي الّذي تناول موضوع التنقيب عن البيانات؛ وذلك لاستيعاب المفاهيم الأساسية ودراسة الأساليب والخوارزميات.
2- زيارة قسم الكبار لدى فرع المكتبة بالدقي وملاحظة كيفية التعامل مع سجلات الإعارة بالنظام الآلي داخل المكتبة؛ ووجد أن لكل مستفيد مشترك في عضوية المكتبة، يوجد له ملف مسجل عليه باسم مستخدم" رقم العضوية" وكلمة مرور من خلال My Account.
3- تجميع تسجيلات ملفات الإعارة من قسم الكبار في مكتبة مصر العامة بالدقي؛ وكان عددها 379,076 تسجيلة، واشتملت على مجموعة من البيانات الآتية: (اسم الكتاب– الموضوع– اللغة)، وذلك في الفترة الزمنية من عام 2011– وحتى ديسمبر 2023.
4- الدراسة والتدريب العملي على برنامج Orange3، وبيان بنيته الهيكلية وآليات تطبيقه في استخراج المعرفة من ملفات بيانات الإعارة للكبار في مكتبة مصر العامة؛ من خلال تنصيب البرنامج، والتطبيق، والاستعانة بالمصادر التعليمية المختلفة، وقد أُختير هذا البرنامج للأسباب الآتية:
   - برنامج مفتوح المصدر يقوم بكل وظائف التنقيب عن البيانات.
   - له واجهة تعامل رسومية إبداعية.
   - يسهل التعامل معه من قِبل غير المتخصصين في علوم الحاسب.
   - يُدعم اللغة العربية.
   - يعرض النتائج النهائية بأكثر من أسلوب للعرض (العرض التفاعلي).
5- إنشاء وتشغيل البرنامج عن طريق رفع مجموعة الملفات عن بيانات الكتب المستعارة بالمكتبة.
6- عرض نتائج التنقيب عن بيانات الإعارة للكبار في مكتبة مصر العامة على متخذي القرار بالمكتبة؛ لتقديم المقترحات حول اكتشاف رؤى جديدة لمجموعات المكتبة، وتقديم الخدمات المعلوماتية التي تتناسب مع احتياجات المستفيدين منها.

7/1 الدراسات السابقة:

أجري مسح شامل للإنتاج الفكري المَنشور حول موضوع الدراسة باللغة العربية والإنجليزية، وذلك باستخدام أدوات بحث الإنتاج الفكري، وباستخدام الكلمات الدالة عن الموضوع؛ ومنها: (التنقيب عن البيانات، وسلوك المستفيدين، والتنقيب عن البيانات **و** المكتبات، والتنقيب عن البيانات **و** خدمات المعلومات، Data Mining, Data Mining(and) Libraries , Data Mining(and)Information Services، وذلك للتعرُّف على الدراسات والجهود الَّتي ترتبط بشكل

مباشِر أو غير مباشر بموضوع الدراسة؛ ووجدت رسالة دكتوراه بعنوان: " استخدام تقنية التنقيب عن البيانات في تطوير المكتبات الرقمية العربية: دراسة تجريبية" عُرض بها 212 دراسة علمية حتى نهاية عام 2011م، تتناول موضوعاتها: استخدام تقنية التنقيب عن البيانات لتطوير خدمات وأنشطة عمليات المكتبات، وتطوير البحث والاسترجاع بالمكتبات الرقمية، ومنعا لتكرار الجهود تستعرض الدراسة الإنتاج الفكري المنشور بعد عام 2011م؛ ولعل منها دراسة (محمد، 2011) والتي تناولت أساسيات التنقيب عن البيانات، وأساسيات الأنطولوجيات، فضلا عن نموذج استرشادي لبناء أنطولوجيات التنقيب عن البيانات؛ واعتمدت الدراسة على المنهج الوصفي التحليلي وأداة تحليل المحتوى لجمع البيانات، وتوصلت الدراسة إلى العديد من النتائج منها: تصميم أنطولوجيات التنقيب عن البيانات عملية معقدة تتطلب خطوات مدروسة لضمان فعاليتها، وتشمل هذه الخطوات صياغة رؤيا عامة لبناء نظام ذكي للتنقيب عن البيانات، ووظائف تكنولوجيا التنقيب عن البيانات، والعلاقات بين أنطولوجية التنقيب عن البيانات ومكونات النظام، وأخيرًا تنفيذ المعرفة ودمجها داخل أنطولوجية التنقيب عن البيانات.

وركز (النشرتي،2012) على رصد توظيف مبادرة الأنطولوجيات في إطار التنقيب عن البيانات والأنطولوجيا للتنقيب على الويب لتحقيق التكامل المعرفي للمصادر على شبكة الويب؛ وذلك بالاعتماد على المنهج الوصفي والمنهج التحليلي، وتوصلت الدراسة إلى عِدة نتائج منها: التحدي الرئيسي للتنقيب على الويب في التعامل مع محتوى الويب الخفي نظرا لعدم الوصول إليه ومن ثم عدم تكشيفه، وأن الأنطولوجيا تساعد على توفير التصنيف الموضوعي للمحتوى المتاح، مما يكفل القدرة الهائلة على استرجاعه، كما تكفل فرص إكساب البرمجيات والتطبيقات القدرة على استنباط حقائق جديدة من واقع المصادر المتاحة، وأوصت الدراسة بأنه يجب تطوير محركات بحث دلالية عربية تتعاظم فيها قدرات البحث والاسترجاع تعتمد على نظم الذكاء الاصطناعي، وتطوير نماذج أنطولوجية بهدف التنقيب عن المعرفة في الملفات الصوتية والفيديو والصور الثابتة، وأخيرا إثراء المحتوى العربي بالأنطولوجيات العامة والمتخصصة.

وعرض (سيد،2016) ماهية التنقيب عن البيانات، وما مهامه وتطبيقاته، وتحليل أدوات التنقيب عن البيانات مفتوحة المصدر وهي:Rapid K, Miner, Weka, Orange ؛ وتقييمها من خلال أربعة معايير للتقييم وهي: الأداء، والوظيفة، والقدرة على الاستخدام، ودعم الأنشطة الثانوية، وتوصلت الدراسة للعديد من النتائج أهمها: هناك ميزة يتصف بها بعض أدوات التنقيب عن البيانات، وتتضح من خلال الاستخدام وهي توفير نموذج السحب والإفلات أثناء عملية التركيب والبناء للتنقيب عن البيانات.

وقدم (Ochilbek, R,2019) طريقة تصنيف للتنبؤ بمعدل استعارة الكتب في المكتبات الأكاديمية من قِبل الطلاب، بناءً على سلوكيات استخدامهم للمكتبة، وشملت البيانات 200 طالب جامعي حول استخدامهم للمكتبة.

وكشفت دراسة (Hu, J.& Tang, J.& Zhang, B.,2019) أن هناك علاقة معينة بين سلوك استعارة القراء والكتب، وبين التخصصات المختلفة، وبين الأنواع المختلفة من القراء واستعارة الكتب، واكتشاف الارتباط بين هذه البيانات يؤدي إلى التخصيص الرشيد للموارد وتحسين استخدام الموارد وتحسين مستوى خدمة المكتبة.

وعرضت (سها علوان،2020) مراجعة علمية للإنتاج الفكري المنشور في موضوع تنقيب البيانات في مجال المكتبات والمعلومات باللغة العربية والإنجليزية خلال الفترة الزمنية من عام 2010 حتى عام 2019، من خلال وصف وتحليل هذا الإنتاج بهدف التعرف على الاتجاهات العددية والنوعية والشكلية والزمنية لهذا الإنتاج للخروج بأهم السمات الموضوعية له، وخرجت الدراسة بمجموعة من النتائج والتوصيات التي حددت في النهاية الاتجاهات البحثية الجديدة والمستقبلية لدراسة موضوع التنقيب عن البيانات في مجال المكتبات والمعلومات، وكيفية تطبيقها في المكتبات ومؤسسات المعلومات.

وقام كل من (Silwattananusarn, T.& Kulkanjanapiban, P.,2020) بدراسة سلوك استخدام المستفيد في المكتبة الأكاديمية ونمط استعارتهم للكتب في مركز خونييغ لونغ لمصادر التعلم بجامعة أمير سونكلا، والتي تؤثر على التحصيل الأكاديمي للمستفيد خلال العام الدراسي 2015–2018؛ وتم استخدام برنامج WEKA كأداة لاستخراج البيانات وتحليلها لتحديد خصائص استعارة كتب المستفيد، ولاكتشاف قواعد الارتباط التي تهم المستفيد، وتحليل العلاقات بين استخدام المكتبة الأكاديمية، وإنجاز طلاب المرحلة الجامعية، وتوصلت الدراسة أن القدرة على تحديد ووصف نمط سلوك المستفيد من المكتبة بشكل واضح، يمكن أن تساعد المكتبة في إدارة الموارد والخدمات بشكل أكثر فعالية.

كما تناول كل من (Ansari, N.& Vakilimofrad, H.& Mansoorizadeh, M.& Amiri, M.R.,2020) تحليل سلوك المستخدم والتنبؤ به في جامعة همدان للعلوم الطبية باستخدام تقنيات استخراج البيانات وخوارزمية k–means، وأظهرت نتائج الدراسة أن تقنيات التنقيب عن البيانات المختلفة قابلة للتطبيق بكفاءة ودقة عالية في تحليل بيانات المكتبات ومؤسسات المعلومات، ويمكن استخدامها للتنبؤ بسلوك المستخدم، وتقديم خدمات فعالة له وإمداده بمصادر المعلومات التي تتوافق مع احتياجاته.

وتناول (Zhang Jing,2021) خوارزمية تطبيق خوارزمية الحاسب الآلي وتكنولوجيا استخراج البيانات الضخمة في خدمة التوصيات الشخصية للمكتبة الذكية، وتحليل تفضيلات القراءة وعادات السلوك لدى القراء، وتقترح ثلاثة مقاييس كمرجع.

وقام كل من (Wei, Machica,2022) بتحليل بيانات سلوك الاستعارة للقارئ باستخدام تقنية التنقيب في البيانات وتحديدا خوارزمية K–Means والعثور على خصائص سلوكيات الإعارة للمستفيدين في نظام الإدارة الآلي للمكتبات بالجامعة؛ وذلك للعثور على اتجاهات القراءة لدى المستفيدين من المكتبة.

واقترح (Pang, Lu,2022) نظام إدارة مكتبات يعتمد على استخراج البيانات وخوارزمية التجميع بالاعتماد على برنامجWEKA، وتوصل الباحث أن الخوارزمية ليست مجرد وسيلة لإدارة أتمتة نظام المكتبة، ولكنها أيضًا وسيلة فعالة لتحقيق تحديث معلومات المكتبة.

وتطرق كل من(Shahnaz Khademizadeh&Zahra Nematollahi & Farshid Danesh,2022) إلى تحليل معاملات تداول الكتب، واكتشاف أنماط إعارة الكتب للمستخدم، وتم تطبيق خوارزميات التصفية التعاونية القائمة على العناصر وطرق شجرة القرار، وتوصلت الدراسة إلى استخلاص قواعد اقتراح الكتب للمستخدمين والتي يمكن الاستفادة منها في التقييم، وتحديد سياسات اقتناء المجموعة، وتخصيص التمويل للمواد، واقتراح أساليب لإلغاء اختيار وتخصيص مساحات الكتب.

كما تناول كل من (حسن& زين العابدين،2023) التعريف بمفهوم عملية التنقيب عن البيانات والكشف عن دلالاتها، وبيان دورها في اكتشاف المعرفة ودراسة أنواع العلاقات التي تنشأ بين البيانات وتجميعها وتصنيفها، كما اقترحا نموذج استرشادي لأداة التنقيب في بيانات المكتبات الجامعية، وتوصلت الدراسة أن تقنية التنقيب عن البيانات تمتلك الأدوات المثلى لتحليل ومعالجة البيانات الضخمة في المكتبات الجامعية، والتي قد تساعد كثيرا في التنبؤ بالقيم المستقبلية، وعلى ضوء ذلك يتم اتخاذ القرارات الصائبة في التحسين من الخدمات في هذا النوع من المكتبات.

وعرض (Qin, 2023) تطبيق تقنيات استخراج البيانات وتأثيرها على خدمات المكتبة؛ بهدف تعزيز كفاءة وفعالية خدمات المكتبة، وتم إنشاء نظام لهذا الغرض يستخدم خوارزمية التجميع المزدوج والتي تم تصميمها لتحديد القيم المتطرفة في البيانات التي يتم تحليلها بعد ذلك، لمعرفة قيمتها المحتملة في توفير اتجاهات جديدة لخدمات المكتبة، وتشير الدراسة أن استخراج البيانات يمكن أن يعزز من تقديم خدمات المكتبات وتطويرها.

- ومن خلال عرض الدراسات السابقة من أدبيات الإنتاج الفكري لوحظ ما يلي:

- نالت تقنية التنقيب عن البيانات الكثير من الاهتمام في الأوساط البحثية على مدى العقد الماضي، في محاولة لتطوير خوارزميات قابلة للتكيف مع الكميات المتزايدة من البيانات؛ وذلك للبحث عن أنماط معرفية ذات معنى.
- اتسمت الدارسات السابقة المنشورة بالتركيز على تطبيق تقنية التنقيب عن البيانات في مجال المحاسبة والبنوك والمجال الطبي والمجال التعليمي؛ ومنعا لتكرار الجهود عُرضت الدراسات التي لها علاقة مباشرة بموضوع الدراسة.
- كانت هناك وفرة في الإنتاج الفكري الأجنبي الذي تناول موضوع التنقيب عن البيانات وتقنياته؛ بينما كانت هناك ندرة في الإنتاج الفكري العربي للموضوع.
- أغلب الدراسات السابقة تعرضت لبرنامج WEKA كأحد برمجيات التنقيب عن البيانات، وندرة من الدراسات تناولت برنامج Orange3 في معالجة النصوص، وبناء نماذج التعلم الآلي، وتحليل المشاعر؛ ولم تتعرض أي من الدراسات إلى دراسة خطوات سير عملWork Flow البرنامج في التنقيب عن البيانات، مما يجعل الدراسة الحالية هي أولى الدراسات–على حد علم الباحثة– التي اعتمدت على تطبيق برنامج Orange3 مفتوح المصدر يقوم بكل وظائف التنقيب عن البيانات ويسهل التعامل معه من قِبل غير المتخصصين.

ثانيا: الإطار النظري:

يستعرض الإطار النظري للدراسة مفاهيم التنقيب عن البيانات، وما أساليبه، وأدواته، والمجالات المرتبطة به، إضافَة إلى تطبيقه في المكتبات ومؤسسات المعلومات.

أولا: التنقيب عن البيانات Data Mining: المفهوم والأساليب والمراحل والأدوات:

1/1/2 النشأة والمفهوم:

ذكر (Martin Brown,2012) أن بداية ظهور إرهاصات الحديث لماهية استخراج البيانات كانت خلال نظرية بايز في عام 1763، واكتشاف تحليل الانحدار في عام 1805، ومن خلال آلة تورينج العالمية (1936)، واكتشاف الشبكات العصبية (1943)، وتطوير قواعد البيانات(السبعينيات)، والخوارزميات الجينية (1975)، واكتشاف المعرفة في قواعد البيانات (1989).

وظهر مصطلح التنقيب عن البيانات فعليا في التسعينيات من القرن العشرين، وتحديدا عام 1996 عندما ظهر كإشارة على هامش اجتماعات اللجنة الدولية المشتركة للذكاء الاصطناعي International Joint Conferences on Artificial Intelligence والمنعقدة في ولاية ميتشجان في الولايات المتحدة (النشرتي، 2012).

وعُرفَ مصطلح التنقيب عن البيانات Data Mining في قاموسODLIS أنه: "عملية استخدام تطبيقات قواعد البيانات لتحديد الأنماط والعلاقات التي لم يتم اكتشافها مسبقًا ضمن مجموعة ضخمة من البيانات" (Reitz, 2004)؛ وبالرجوع للإنتاج الفكري المنشور وجدت العديد من التعريفات التي تناولت مفهوم التنقيب عن البيانات Data Mining؛ ونذكر منها ما يلي:

- العملية التي تعتمد على استخدام الذكاء الاصطناعي والأساليب الإحصائية في استخراج المعلومات المفيدة من قواعد البيانات الضخمة (هاني، 2018).
- استخدام التعلم الآلي والتحليل الإحصائي للكشف عن الأنماط والمعلومات القيمة الأخرى من مجموعات البيانات الكبيرة.(Martínez & Chaparro, 2016)

- الأساليب والأدوات المنهجية لاستكشاف المعرفة الخفية في قواعد بيانات المنظمات؛ ويمكن النظر إليه كشكل ضروري من أشكال المعرفة لحل المشكلات المحددة في مجال معين، كما يمكن بناء نموذج قرار للتنبؤ وتصنيف المشاكل المحتملة (محمد، 2015).
- مجموعة من التقنيات التي تعتمد على الحاسوب والمصممة لاستخراج المعلومات من أحجام كبيرة من البيانات الخفية أو غير المتوقعة، وتعرف عملية التنقيب على البيانات أحيانا باكتشاف المعرفة في قواعد البيانات (Rebbapragada &Semple, 2010)

وباستعراض عددٍ من التعريفات التي تناولت مفهوم التنقيب عن البياناتData Mining؛ وُجد أن هناك العديد من المصطلحات التي تدل على مفهوم التنقيب عن البيانات في الإنتاج الفكري المنشور مثل: " تنقيب البيانات– التنقيب عن البيانات– التنقيب في البيانات– استخراج البيانات– التنقيب عن المعلومات– اكتشاف المعرفة".

ويمكننا القول: إن التنقيب عن البيانات يُعد أسلوبا متقدما من أساليب تحليل البيانات، يقوم بالبحث والاستكشاف من أجل الحصول على الأنماط غير المعروفة، والعلاقات المهمة بين مجموعات البيانات المختلفة عن طريق توظيف أساليب الذكاء الاصطناعي والإحصاء؛ من أجل الحصول على المعلومات التي تساعد متخذي القرار في بناء التنبؤات المستقبلية واستكشاف السلوك والاتجاهات، مما يساعد على اتخاذ القرارات السليمة في الوقت المناسب.

ومن خلال ما سبق، يمكننا القول بأن تقنية التنقيب عن البيانات تعتمد على ما يلي:(سيد2016)

- البيانات Data : هي عبارة عن الحقائق والأرقام والنصوص في صورتها الأولية والتي لا يمكن الاعتماد عليها في التحليل.
- المعلومات Information: النماذج والعلاقات بين تلك البيانات ومعالجتها وتمثيلها وتقديمها بصورة مناسبة والتي تشكل معلومات مفيدة.
- المعرفة Knowledge: المعلومات السابقة يمكن أن تحول إلى معرفة، ويمكن بناء التنبؤات التي تساعد في اتخاذ القرارات.
- مستودعات البيانات: تستخدم في التحليلات الزمنية واكتشاف المعرفة واتخاذ القرارات.

وذكر (Bin, C,2013) أن هناك مجموعة من الشروط التي يجب أن تتوافر في البيانات المراد تطبيق التنقيب عن البيانات في المؤسسات المختلفة؛ وهى: (دقة وتكامل وسلامة البيانات، والمصداقية، وسهولة الحصول عليها والوصول إليها في أي وقت).

2/1/2 الخصائص والسمات:

يوضح (سمير، 2016) أن أساليب التنقيب في البيانات تتميز بالعديد من الخصائص والسمات؛ منها:

1. الاكتشاف الآلي لأنماط غير معروفة: تعمل أساليب التنقيب في البيانات على اكتشاف أنماط البيانات آليا، ويساعد ذلك على اكتشاف الغش والأخطاء.
2. تستطيع أساليب التنقيب في البيانات التعامل مع حجم كبير من البيانات: ومن ثَم تعتبر أحد المزايا المهمة عند أداء عمليات المراجعة.
3. القدرة على التعلم: إن تلك الأساليب تكتسب خبرة من أخطائها السابقة وتتحسن جودة هذه النماذج تلقائيا.
4. وذكر (Martin Brown,2012) أن تطبيق تقنية التنقيب عن البيانات داخل المؤسسات المختلفة؛ يهدف إلى تحقيق ما يلي:
   - استكشاف البيانات عن العملاء، وفهمها بشكل أفضل.
   - الكشف عن الأنماط والاتجاهات الكامنة في البيانات، عن طريق تحليل البيانات الضخمة، وتحديد الاتجاهات المستقبلية المحتملة والتوقعات بناءً على البيانات التاريخية التي سبق التوصل إليها.

- استخدام البيانات المتاحة للاستنتاج والتنبؤ بالأحداث المستقبلية للعملاء والمؤسسة.
- توفير الدعم لاتخاذ القرارات، من خلال استخدام الأنماط والنماذج لمساعدة المديرين وصناع القرار في فهم المشكلات والتحديات التي تواجه المؤسسة.
- اكتشاف الاحتياجات الناشئة والفرص الجديدة، واكتشاف الاتجاهات الصاعدة والفرص الناشئة في السوق، واستغلالها في الحصول على ميزات تنافسية بين المؤسسات وبعضها.
- تحسين الأداء في العمليات، والعمليات التشغيلية وتحسينها لتحقيق أفضل كفاءة وفاعلية.

3/1/2 أنواع التنقيب عن البيانات:

يستهدف كل نوع من أنواع التنقيب عن البيانات نوعًا محددًا من البيانات وأهدافًا تحليلية مختلفة؛ ومن بين هذه الأنواع: (Peker& Osman,2018)

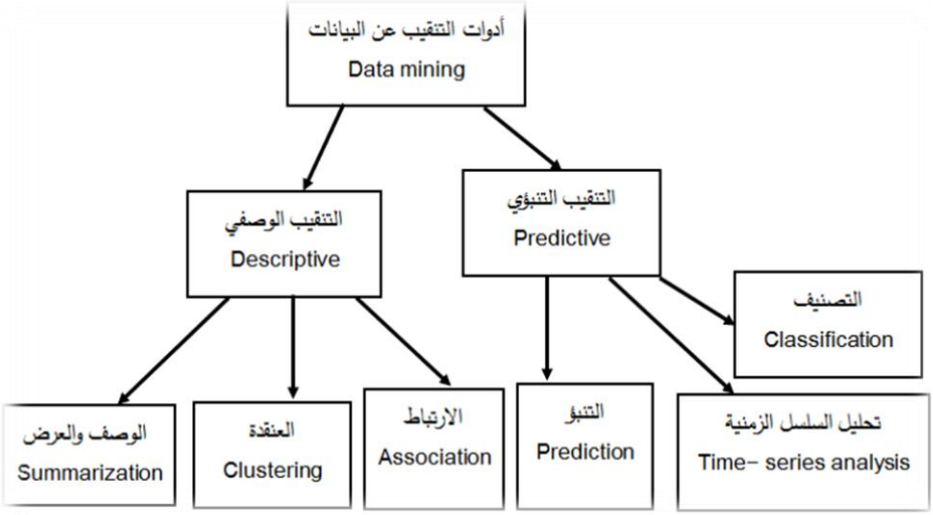
- التنقيب في العمليات: يهدف هذا النوع إلى استخلاص المعرفة من سجلات الأحداث المتاحة في أنظمة المعلومات؛ بهدف فهم وتحسين علاقات المؤسسة بمورّديها.
- التنقيب في النصوص: يهتم هذا النوع من التنقيب باستخراج المعرفة من النصوص الرقمية مثل: المقالات والمستندات ورسائل البريد الإلكتروني، يتم استخدامه على نطاق واسع في تحليل مشاعر العملاء وتصنيف المحتوى.
- التنقيب التنبؤي: يستخدم هذا النوع من التنقيب ذكاء الأعمال للتنبؤ بالتوجهات، ويساعد قادة الأعمال في دراسة تأثير قراراتهم على مستقبل الشركة واتخاذ خيارات فعّالة.

4/1/2 أساليب التنقيب عن البيانات وخوارزمياته:

يذكر (فايز، 2016) أن عملية التنقيب عن البيانات تتضمن عمليتين أساسيتين متمثلتين في التنقيب الوصفي والتنقيب التنبؤي، وفي التنقيب الوصفي يعتمد تحليل البيانات على إعادة تنظيم البيانات والتنقيب فيها لاستخراج المعلومات منها، أما التنقيب التنبؤي فيعتمد تحليل البيانات على استخدام معلومات سابقة لتوقع ما سيحدث في المستقبل.

ولكل عملية من عمليات التنقيب عن البيانات الأدوات والخوارزميات التي تعمل بها؛ تُعد تحليل السلاسل الزمنية والتصنيف والتنبؤ أدوات العمل للتنقيب التنبؤي، أما الارتباط والعنقدة والوصف والعرض فتندرج ضمن أدوات التنقيب الوصفي عن البيانات.

* وفيما يلي عرض لأنواع التنقيب عن البيانات والخوارزميات المستخدمة في كل نوع؛ كما يوضحه الشكل رقم (1):



شكل رقم (1) أساليب وخوارزميات التنقيب عن البيانات.

المصدر:.https://agri-research-journal.net/sjar/wp-content/uploads/2020/06/v7n3p13.pdf

ويمكننا عرض الخوارزميات الأكثر استخداما في التنقيب عن البيانات؛ كما يلي: (Han, Jiawei & Kamber, Micheline& Pei, Jian,2011)

- **خوارزميات التصنيف:** يتم تصنيف البيانات إلى مجموعات فرعية بناءً على خصائصها.
- **خوارزميات التجميع:** تقوم بتجميع البيانات المتشابهة التي تشترك في الخصائص المشتركة في مجموعات فرعية أو مجموعات.
- **خوارزميات التنبؤ:** تتنبأ بالقيم المستقبلية للبيانات بناءً على البيانات المُتاحة.
- **خوارزمية الارتباط:** أسلوب يوضح الارتباط بين عنصرين أو أكثر، غالبًا من النوع نفسه لتحديد الأنماط.
- **خوارزمية تحليل التسلسل:** أسلوب للبحث عن الأنماط التي تؤدي فيها مجموعة معينة من الأحداث أو القيم إلى أحداث أو قيم لاحقة.
- **خوارزمية شجرة القرار:** تقنية تستخدم تحليلات التصنيف أو الانحدار لتصنيف النتائج المحتملة أو التنبؤ بها بناءً على مجموعة من القرارات؛ ويستخدم هذا النمط تصورًا يشبه الشجرة لتمثيل النتائج المحتملة لهذه القرارات.
- **خوارزمية KNN:** تُصنف البيانات في هذه الخوارزمية بناءً على قربها وارتباطها بالبيانات الأخرى المتاحة؛ وتفترض وجود نقاط بيانات متشابهة بالقرب من بعضها البعض، ونتيجة لذلك تسعى إلى حساب المسافة بين نقاط البيانات أو المتوسط.

5/1/2 مراحل التنقيب عن البيانات:

يرى(الفارسي،2022) أن مصطلح اكتشاف المعرفة في قواعد البيانات "KDD" Knowledge Discovery In Databases بشكل عام يشير إلى العملية الشاملة لاستخراج المعرفة المفيدة من البيانات، والتنقيب في البيانات خطوة في هذه العملية والتي تتضمن العديد من المراحل لن تتوقف عند تجميع البيانات وإدارتها؛ بل تمتد لتشمل التحليل والتنبؤ بما يحدث مستقبلا؛ وتتمثل هذه المراحل في:

1. اكتشاف البيانات Data Discovery: مرحلة تجميع البيانات، وفيها يتم كشف وتحديد وتوصيف البيانات المتاحة، وفهمها بشكل أفضل.
2. تصفية البيانات وتنقيتها Data Cleaning: مرحلة حذف البيانات غير المهمة والمتضاربة وغير المتناسقة.
3. تكامل البيانات Data Integration: مرحلة تجميع البيانات المتشابهة، وذات الصلة من مصادر البيانات المتعددة ودمجها معا.
4. اختيار البياناتData Selection: مرحلة تحديد واسترجاع البيانات الملائمة من مجموعة البيانات.
5. تحويل البيانات Data Transformation: مرحلة تحويل البيانات إلى نماذج مخصصة ملائمة لإجراءات البحث والاسترجاع.
6. التنقيب عن البيانات Data Mining: مرحلة الكشف عن الأنماط والاتجاهات الكامنة في البيانات باستخدام طرق ذكية لاستخراج الأنماط المفيدة من البيانات، عن طريق تحليل البيانات الضخمة لتحديد الاتجاهات المستقبلية المحتملة والتوقعات.
7. تقييم النمط Pattern Evaluation: مرحلة تحديد الأنماط المهمة والتي تمثل المعرفة، عن طريق استخدام البيانات المتاحة للاستنتاج والتنبؤ، وتوليد نماذج وتوقعات تساعد في اتخاذ القرارات والتخطيط للمستقبل.
8. تمثيل المعرفة Knowledge Implementation: وهي المرحلة الأخيرة من مراحل اكتشاف المعرفة في قواعد البيانات والتي يراها المستفيد وتساعده على فهم وتفسير نتائج استخراج البيانات.

6/1/2 أدوات التنقيب عن البيانات:

هناك العديد من الأدوات المستخدمة في التنقيب عن البيانات تختلف في الخصائص والمميزات التي تقدمها، سواء أكانت مفتوحة المصدر أو بمقابل مادي؛ ويوضحها الشكل رقم (2) والجدول رقم (1):

شكل رقم(2) أدوات التنقيب عن البيانات.

المصدر: https://data-flair.training/blogs/data-mining-tools-techniques

وبالبحث والاطلاع على الإنتاج الفكري؛ تم حصر 16 أداة للتنقيب عن البيانات بعضها متاح بمقابل مادي وبعضها مجانية؛ كما هو موضح في الجدول رقم (1): (Simplilearn,2024)

جدول رقم(1) أدوات التنقيب عن البيانات

م	اسم الأداة	التعريف بها
.1	**Apache Mahout**	طورتها مؤسسة Apache؛ وتُعد أحد أفضل أدوات التنقيب عن البيانات مفتوحة المصدر في السوق خصيصا للباحثين والمهنيين الذين يعتزمون تنفيذ الخوارزميات الخاصة بهم، وإنشاء تطبيقات قابلة للتطوير باستخدام ممارسات تعلم الآلة، تم بناؤها بلغة البرمجة JAVA ويمكنها التعامل مع مجموعات كبيرة من البيانات وتفضلها شركات مثل LinkedIn وYahoo.
.2	**SAS Enterprise Miner**	تم تطويرها بواسطة معهد نظام التحليل الإحصائيSAS Institute؛ توفر الأداة واجهة بديهية تساعد في بناء نموذج تحليلي أسرع، وتدعم الخوارزميات المختلفة لتساعد المحترفين على إدارة البيانات بشكل أفضل؛ وذلك عن طريق تحويل أجزاء كبيرة من البيانات إلى رؤى قيمة، وتُستخدم في العديد من المجالات مثل: الأعمال التجارية والصناعة والرعاية الصحية والتسويق والمالية والعلوم الاجتماعية.
.3	**Orange**	عبارة عن مجموعة أدوات شاملة مفتوحة المصدر للتنقيب عن البيانات والتعلم الآلي وتصوير البيانات؛ توفر الأداة واجهة سهلة الاستخدام مع ميزات رسومية إضافية تجعل تصور البيانات أكثر تفاعلية، كما تدعم البرمجة المرئية، ويمكن لغير الخبراء في المجال إنشاء نماذج ببساطة باستخدام ميزات واجهة السحب والإفلات، وفي الوقت نفسه، يمكن للمهنيين المهرة الاعتماد على لغة برمجة Python لتطوير النماذج.
.4	**Dundas BI**	تُعد واحدة من الأدوات الأكثر شمولاً والمستخدمة لتوليد رؤى سريعة حول تبسيط معالجة البيانات وتحليلها، كما تتميز بإمكانية الوصول إلى البيانات من أجهزة متعددة، كما يمكنه تحليل البيانات متعددة الأبعاد وإصدار التقارير.
.5	**Oracle Data Mining**	تعد تلك الأداة امتدادًا لـــ Oracle SQL Developer Advanced Analytics؛ صُممت لعلماء ومحللي البيانات بمقابل مادي لتوفير مجموعة من الوظائف والأدوات لبناء نماذج التعلم الآلي وتقييمها وتحليلها داخل بيئة قاعدة البيانات، وتنشئ نماذج تنبؤية وتتضمن خوارزميات متعددة ضرورية لمهام مثل: التصنيف والانحدار والتنبؤ. يمكن للمستخدمين دمج هذه الخوارزميات في استعلامات SQL الخاصة بهم، مما يسمح ببناء النماذج وعمليات التقييم بكفاءة داخل واجهة Oracle SQL Developer المألوفة.

م	اسم الأداة	التعريف بها
.6	**RapidMiner**	هي منصة مفتوحة المصدر تهدف إلى توفير بيئة تطوير مرئية وسهلة الاستخدام للمحللين والعلماء في مجال البيانات وتدعم العديد من الخوارزميات الأساسية للتعلم الآلي والتعلم العميق، واستخراج النصوص والتحليلات التنبؤية، كما توفر مجموعة من النماذج المُعدة مسبقًا والتي تساعد غير الخبراء على تطوير سير العمل دون الحاجة إلى برمجة صريحة في سيناريوهات محددة، يتم دعم الأداة أيضًا من قِبل مستخدمي المجتمع النشطين المتوفرين دائمًا للمساعدة، ويمكن للمطورين الاستفادة من فوائد R وPython لبناء نماذج تحليلية تتيح تصور الاتجاه والنمط والتصور الخارجي.
.7	**IBM SPSS Modeler**	أداة مملوكة في الأصل لشركة SPSS Inc ثم طُورت بعد ذلك بواسطة شركة IBM؛ تُتاح بمقابل مادي وتوفر التنقيب عن البيانات في نسختين IBM SPSS Modeler Professional -IBM SPSS Modeler Premium. تتميز الأداة بواجهة رسومية سهلة الاستخدام تتضمن ميزات إضافية لتحليلات الكيانات والنصوص، وتسمح للمستخدمين بتحليل البيانات واستكشاف الأنماط وبناء نماذج تنبؤية بشكل سهل وفعال، كما تعتمد على سحب وإفلات العناصر والعمليات المختلفة لبناء سلاسل تحليل البيانات، وتتيح استيراد كميات كبيرة من البيانات من عدة مصادر متباينة للكشف عن أنماط واتجاهات البيانات المخفية.
.8	**WEKA**	أداة ML مفتوحة المصدر متاحة للتنزيل مجانًا باستخدام رخصة GNU العامة ومكتوبة بلغة JavaScript؛ تم تطويره من قبل باحثين في جامعة وايكاتو في نيوزيلندا، توفر الأداة واجهة سهلة الاستخدام مع ميزات أنه يحتوي على خوارزميات وأدوات التصور التي تدعم التعلم الآلي، كما تسمح للمستخدمين ببناء النماذج دون كتابة التعليمات البرمجية، وهذا يتطلب معرفة جيدة بالخوارزميات المستخدمة لمثل هذه الأغراض حتى يتم اختيار الخوارزمية المناسبة بشكل صحيح، تم تصميم أداة WEKA في البداية لاستكشاف القطاع الزراعي، واليوم يتم استخدامه على نطاق واسع من قِبل الباحثين والعلماء لاستكشاف القطاع الأكاديمي.
.9	**Konstanz"KNIME Information " Miner**	أداة مفتوحة المصدر، تم تطويرها بواسطة شركة KNIME AG، وتوفر مجموعة واسعة من الوظائف والأدوات للتنقيب عن البيانات وبناء سلاسل التحليل، فهي توفر مجموعة واسعة من الأدوات للتحليل الإحصائي، والتعلم الآلي، واستخراج النص، وتحليل الصور. وتتميز بواجهة رسومية سهلة الاستخدام تعتمد على سحب وإفلات العناصر والعمليات المختلفة لبناء سلاسل التحليل، وبقدرتها على دمج مكونات مختلفة للتعلم الآلي واستخراج البيانات من خلال مفهوم خطوط أنابيب البيانات المعيارية.
.10	**Monkey Learn**	تستخدم تقنيات التعلم الآلي لتحويل النصوص والبيانات غير المنتظمة إلى بيانات منظمة ومفهومة، ويمكن أن تساعد المنصة الشركات في استخلاص الرؤى من المصادر النصية مثل: منشورات ووسائل التواصل الاجتماعي ومراجعات العملاء والمقالات النصية، وتتميز هذه المنصة بأنها: سهلة الاستخدام، وإمكانية التكامل مع المنصات الأخرى، كما يمكن التعامل مع كميات كبيرة من البيانات.
.11	**H2O**	منصة مفتوحة المصدر وقابلة للتطوير للتعلم الآلي والتحليلات التنبؤية؛ فهي تتيح إنشاء نماذج التعلم الآلي على مجموعات بيانات ضخمة، إلى جانب النشر المباشر لتلك النماذج داخل بيئة المؤسسة، وعلى الرغم أن قاعدة التعليمات البرمجية الأساسية لـ H2O هي Java، إلا أنها توفر إمكانية الوصول من خلال واجهات برمجة التطبيقات في Python وR وScala؛ والتي تلبي احتياجات العديد من المطورين وعلماء البيانات، وتتميز أداة H2O بأنها مصممة للقياس أفقيًا للتعامل مع مجموعات البيانات الكبيرة، كما تدعم معظم خوارزميات التعلم الآلي الرئيسية.
.12	**Teradata**	تُعرف باسم Teradata Database؛ وتناسب المؤسسات التي تعتمد على إعدادات النشر السحابية المتعددة ومتاحة للنشر على أية منصة سحابية عامة مثل،AWS ، Azure Google، تتميز بمستودع بيانات على مستوى المؤسسات لإدارة البيانات واستخراج البيانات بشكل سهل، والحصول على رؤى حول البيانات المهمة للأعمال المتعلقة بتفضيلات العملاء، وتحديد موضع المنتج، والمبيعات.

م	اسم الأداة	التعريف بها
.13	**Data Melt**	هي أداة مفتوحة المصدر ومعروفة أيضًا باسم DMelt، وهى أداة مساعدة متعددة المنصات مكتوبة بلغة JAVA، ويمكن تشغيلها على أي نظام تشغيل وتتوافق مع"JVM" Java Virtual Machine، وتُستخدم بشكل أساسي في التحليل الإحصائي وتصور البيانات في المجالات التي تتعامل مع كميات هائلة من البيانات، مثل: الأسواق المالية.
.14	**Rattle**	أداة مفتوحة المصدر تعتمد على واجهة المستخدم الرسومية، وتعمل على تعزيز قدرات الحوسبة الإحصائية القوية للغة البرمجة R وتمكن للمستخدمين من إنشاء تعليمات برمجية، وتتميز بقيامها بوظائف واسعة النطاق لاستخراج البيانات، وواجهة مستخدم مصممة بشكل جيد، كما تمنح الطرف الآخر إمكانية مراجعة الكود واستخدامه لأغراض عديدة، وتوسيع الكود دون أي قيود.
.15	**Amazon Sage Maker**	منصة برمجيات؛ تساعد المحللين والمطورين في إعداد نماذج تعلم آلة عالية الجودة وإنشائها وتدريبها ونشرها، وتضم مجموعةً من الأدوات لإجراء عملية التنقيب في البيانات مثل: (Amazon Sage Maker Data Wrangler, Amazon Sage Maker Studio, Amazon Sage Maker Debugger).
.16	**Sisense**	أداة للتنقيب عن البيانات تسمح لمحللي البيانات بدمج البيانات من مصادر متعددة لتطوير المستودع المؤسسي؛ تشتمل على عناصر واجهة مستخدم بالإضافة إلى ميزات السحب والإفلات والتي تعمل على تبسيط عملية تحسين البيانات وتحليلها، كما يمكن للمستخدمين تحديد عناصر واجهة مستخدم مختلفة لإنشاء التقارير بسرعة في مجموعة متنوعة من التنسيقات، بما في ذلك المخططات الخطية والرسوم البيانية الشريطية والمخططات الدائرية.

وذكرت (Alexandra Twin,2024) أن هناك بعض العوامل التي بناءً عليها يتم المفاضلة بين أدوات التنقيب عن البيانات؛ ومنها:

- **كمية البيانات:** لابد أن يتم المفاضلة بين أدوات التنقيب عن البيانات المدفوعة ومفتوحة المصدر بناءً على حجم بيانات المؤسسة.
- **الموارد البشرية:** يعتمد استخدام أدوات التنقيب عن البيانات أيضًا بشكل كبير على الكوادر البشرية المؤهلة لدى المؤسسة.
- **النتائج:** يتوقف اختيار أدوات التنقيب عن البيانات على حجم النتائج التي ترغب بها المؤسسة، ومدى التنبؤ بالنتائج المستقبلية.
- **الدعم والتحديث:** يجب على أداة التنقيب عن البيانات أن توفر دعمًا لمستخدميها باستمرار، ويقوم بترقية منتجاته بانتظام فيما يتعلق باحتياجات العمل المتغيرة.
- **واجهة مستخدم رسومية:** لابد وأن تحتوي أداة التنقيب عن البيانات على واجهة مستخدم سهلة الاستخدام وخالية من الرموز.
- **سهولة الاستخدام:** يفضل أن تكون أداة التنقيب عن البيانات سهلة الاستخدام.
- **دعم أنظمة السحابة:** يفضل أن تتوافق أداة التنقيب عن البيانات مع أنظمة التخزين السحابية؛ بناءً على حجم مؤسستك، من الممكن أن تضطر إلى التعامل مع مصادر ووجهات سحابية.

7/1/2 تطبيقات التنقيب عن البيانات: (Kaderye, Golam& Arif, Ahsan& Kundu, Ronjon,2024)

من خلال مراجعة الإنتاج الفكري المنشور في موضوع التنقيب عن البيانات؛ تبين أن هناك العديد من المجالات التي يمكن تطبيق تقنيات التنقيب عن البيانات فيها، ومنها:

- **التجارة والإحصاء:** من خلال تحليل سلوكيات العملاء، وتطوير الحملات تسويقية، وتقييم المخاطر المالية، وكشف الاحتيالات في القوائم المالية.

- **الطب:** من خلال تحليل بيانات المرضى لتشخيص الأمراض، وتطوير الأدوية لوضع نظام علاجي متكامل.
- **التخطيط الإستراتيجي والتنمية:** من خلال تحسين عمليات الإنتاج، والتنبؤ بالأعطال، ووضع الخطط المستقبلية، ودعم اتخاذ القرارات في المؤسسات المهنية.
- **التعليم:** من خلال تحليل البيانات العلمية، واكتشاف المعرفة الجديدة، والخروج بمؤشرات عامة حول الأداء الأكاديمي لدعم السياسات التعليمية لدى متخذي القرار في المؤسسات التعليمية.
- **تكنولوجيا المعلومات:** من خلال تطبيق التقنيات الحديثة في طرق معالجة النصوص، وبناء نماذج التعلم الآلي، وتحليل المشاعر.

ويرى (سيد، 2016) أنه يمكن الاستفادة من تطبيق تقنية التنقيب عن البيانات في المكتبات ومؤسسات المعلومات في تحسين جودة العملية الوظيفية الداخلية للمكتبات، وتحسين الخدمات المعلوماتية المقدمة؛ وذلك من خلال:

- **إدارة مقتنيات المكتبة:** من خلال الكشف عن اتجاهات المستفيدين، وتوضيح أوجه القصور والزيادة في مقتنيات المكتبة مع إعادة تقسم مجموعات المكتبة، واستنباط رؤى جديدة للمقتنيات يلائم احتياجات المستفيدين منها.
- **دراسات المستفيدين:** من خلال الكشف عن آراء المستفيدين واتجاهات الاستعارة والبحث الخاصة بهم تجاه مقتنيات المكتبة، ووضع كل هذه البيانات في قاعدة بيانات واحدة يمكن من خلالها استخراج معلومات تفيد متخذي القرار عند اقتناء مصادر المعلومات، لكي تلبي احتياجاتهم الفعلية واتجاهاتهم البحثية.
- **خدمات المعلومات:** من خلال تقييم سلوك المستفيدين، وإعادة تقسيم المقتنيات وتكاملها يمكن مساعدة المسؤولين في تقييم الخدمات الفعلية المقدمة والتخطيط الصحيح لتقديم خدمات جديدة لتلبية احتياجات المستفيدين من المكتبة.
- **موارد المكتبة:** فمن خلال إعادة الهيكلة الجديدة لإدارة مقتنيات المكتبة، يصبح هناك بعض الترشيد في أوجه صرف الميزانية وتوفيرها لجوانب جديدة أكثر إفادة.

ثالثا: الإطار التطبيقي:

يستعرض الإطار التطبيقي للدراسة مراحل تطبيق برنامج Orange3 كأحد البرمجيات مفتوحة المصدر التي تستخدم في التنقيب عن البيانات، والمُتاح عبر موقعه الإلكتروني على شبكة الإنترنت. https://orangedatamining.com/

1/3 مراحل تطبيق برنامج Orange3 للتنقيب عن البيانات في تحسين مستوى خدمة الإعارة للكبار في مكتبة مصر العامة:

تعمل مكتبات مصر العامة بكل فروعها على اقتناء أحدث ما ينتجه الفكر الإنساني في العلوم والآداب؛ كما تخدم فئات متعددة من المستفيدين بأعمار مختلفة، فتقدم لهم العديد من المواد الثقافية والتعليمية المتاحة وتنميتها وتحديثها بما يفي برغبات واحتياجات المستفيدين منها؛ أدى ذلك إلى تضخم مصادر المعلومات بها وظهور الحاجة إلى تطوير أساليب أكثر قدرة على إجراء عمليات البحث والتنقيب عن البيانات كتقنية تهدف إلى استنتاج المعرفة من الكميات الهائلة من البيانات، وإتاحة المعلومات بالشكل الذي يعزز من استخدامها.

لقد ظهر مع التنقيب عن البيانات أدوات وبرمجيات تساعد في عملية التنقيب عن هذا الكم الهائل والمتزايد من البيانات للوصول إلى المعرفة في قواعد البيانات المختلفة، وتساعد هذه الأدوات العمل على معظم التخصصات العلمية ومنها علوم المكتبات والمعلومات؛ وفي ضوء الخصائص والمميزات التي تتمتَّع بها تلك الأدوات، تم اللجوء إلى عرض

وتحليل برنامج Orange3 والذي يُعد من أكثر الأدوات شيوعًا في مجال التنقيب عن البيانات وتحليليها، لاستخدامه في تحليل بيانات الإعارة للكبار بمكتبة مصر العامة بفرعها الرئيسي بالدقي.

وجدير بالذكر، برنامج Orange3 يُعد بمثابة بيئة برمجية شاملة مفتوحة المصدر يمكن استخدامه تحت نمط الرخصة العامة (GNO)؛ تم تطويره في مختبر المعلوماتية الحيوية بكلية الحاسوب والمعلومات بجامعة ليوبليانا سلوفينيا ليستخدم في إنشاء وتصميم سير العمل workflows للتنقيب في البيانات، ويمكن تطبيقه بشكل فعال في المؤسسات المختلفة كالمكتبات ومؤسسات المعلومات (Berry,2004)

2/3 مميزات برنامج Orange3:

- برنامج مفتوح المصدر يقوم بكل وظائف التنقيب عن البيانات.
- له واجهة تعامل رسومية إبداعية.
- يسهل التعامل معه من قِبل غير المتخصصين في علوم الحاسب دون الحاجة إلى تعلم لغات برمجة.
- يُدعم اللغة العربية.
- يوفر مجموعةً واسعةً من أدوات التحليل المُتقدمة التي تدعم عمليات البحث في البيانات المهيكلة وغير المهيكلة.
- عرض وتنظيم مجموعات كبيرة من البيانات الخاصة أو العامة بسهولة واستخراج الكيانات والعلاقات منها.
- التكامل مع أدوات أخرى مثل Python لزيادة قدراته التحليلية.
- يعرض النتائج النهائية بأكثر من أسلوب للعرض وبطرقٍ بصريةٍ سهلة الفهم (العرض التفاعلي).

3/3 مجالات تطبيق برنامج Orange3:

يمكن تطبيق تقنية التنقيب عن البيانات باستخدام برنامج Orange3 في العديد من المجالات؛ ومنها:

- التسويق: من خلال تحليل سلوكيات العملاء وتطوير الحملات التسويقية.
- التمويل: في اكتشاف الاحتيال وتقييم المخاطر المالية.
- الرعاية الصحية: في تحليل بيانات المرضى لتشخيص الأمراض وتطوير علاجات جديدة.
- التصنيع: لتحسين عمليات الإنتاج وتنبؤ الأعطال.
- العلوم: لتحليل البيانات العلمية واكتشاف المعرفة الجديدة.

4/3 تطبيق برنامج Orange3 على بيانات الإعارة للكبار في مكتبة مصر العامة:

يوضح الشكل رقم (3)؛ واجهة عمل برنامج Orange3، فمن خلال الأداة يمكنك استيعاب وتحميل البيانات وخوارزميات التعلم الآلي الأساسية والمرئيات التفاعلية؛ كما يتم تجميع أدوات العمل داخل البرنامج في فئات وفقًا لوظيفتها.

☒ يتم العمل على البرنامج من خلال قائمة الأدوات Widget داخل البرنامج؛ سواء بالضغط عليها من القائمة على اليسار أو بالضغط على زر الماوس الأيمن في وسط الصفحة، وكل أداة تقوم بوظيفة معينة مثل: (قراءة البيانات، وعرض جداول البيانات، ورسم البيانات، وبناء نموذج...) يتم وصفها أسفل القائمة بحيث يتم استخدام أي من هذه الوظائف على حسب الحاجة.

شكل رقم (3) تسجيل الدخول على برنامج Orange3

المصدر: https://orangedatamining.com/getting-started/

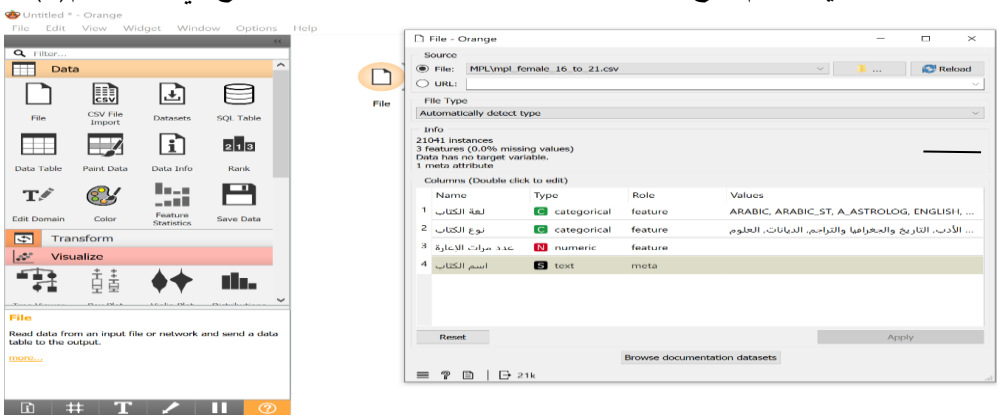
☒ يتم عرض آليات تطبيق البرنامج في استخراج المعرفة من ملفات بيانات الإعارة للكبار في مكتبة مصر العامة؛ وذلك من خلال تطبيق خطوات التنقيب عن البيانات في سير العمل على البرنامج.

- **الخطوة الأولى:** وهي (اكتشاف البيانات وفهمها):

– يتم فتح أداة File؛ كما يظهر في الشكل رقم (4)، ونجد أن هناك عِدة اختيارات لقراءة مجموعة البيانات للبرنامج؛ منها: الضغط على File إذا كانت مجموعة البيانات عبارة عن ملف، أو الضغط على ملف CVS، أو الضغط علىData Set لتحميل الملفات من مستودع بيانات رقمي أو تحميل الـ Data Set من SQL.

شكل رقم (4) التعرف على البيانات وفهمها

– يتم استدعاء مجموعة البيانات Dataset المراد التنقيب بها؛ بالضغط على أداة file لتظهر الملفات التي نريد عمل التنقيب عن البيانات فيها، ويظهر كم ونوع البيانات الموجودة داخل الملف؛ كما هو موضح في الشكل رقم(5):

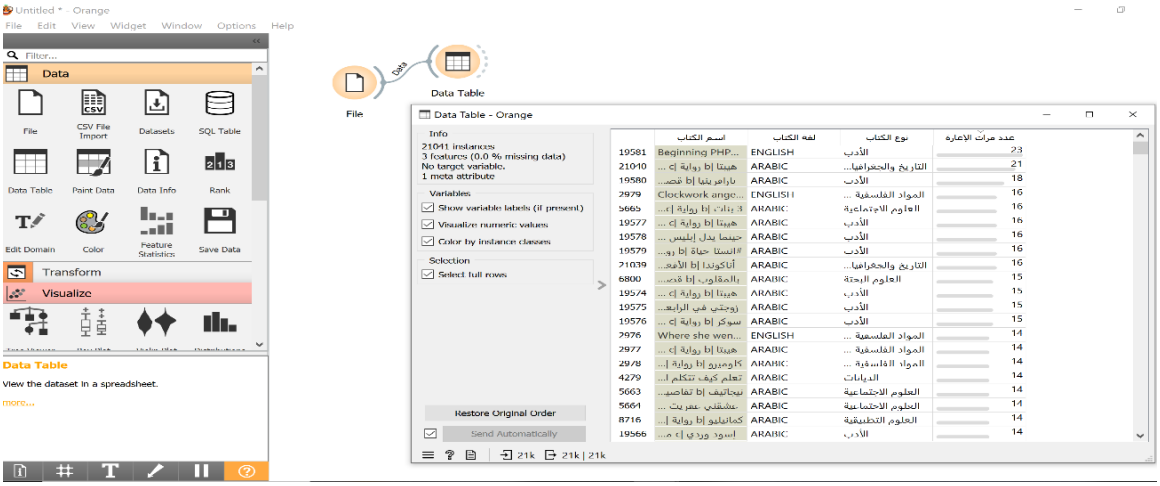


شكل رقم (5) استدعاء مجموعة البيانات Data Set

- وبمجرد الضغط على Apply يتم تحميل مجموعة البيانات داخل البرنامج.

- **الخطوة الثانية:** هي (تصفية البيانات وتنقيتها):

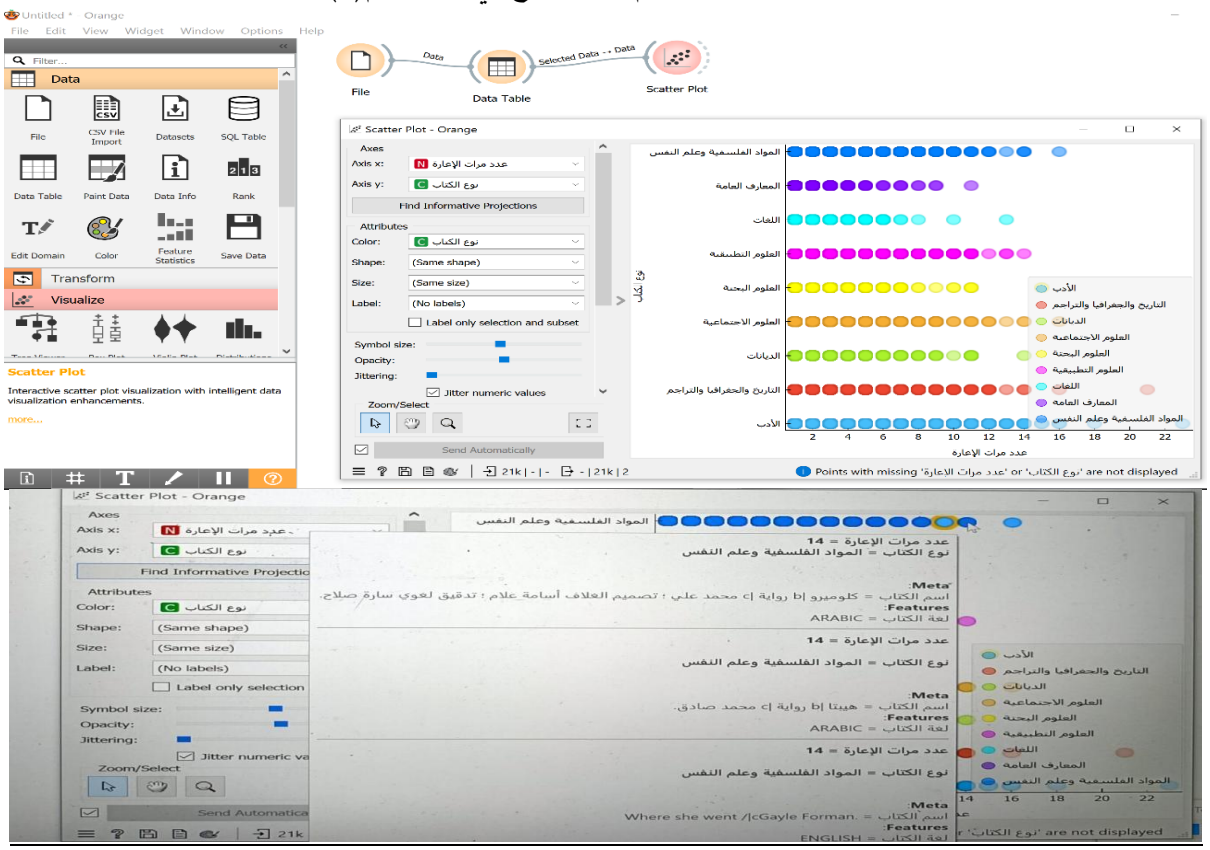
- يتم تنظيف البيانات وعرضها على واجهة البرنامج بمجرد الضغط على زر الماوس الأيمن واختيار أداة Data table لعرض كل الـ Data set values، كما هو موضح في الشكل رقم (6):



شكل رقم (6) عرض البيانات في جداول البيانات.

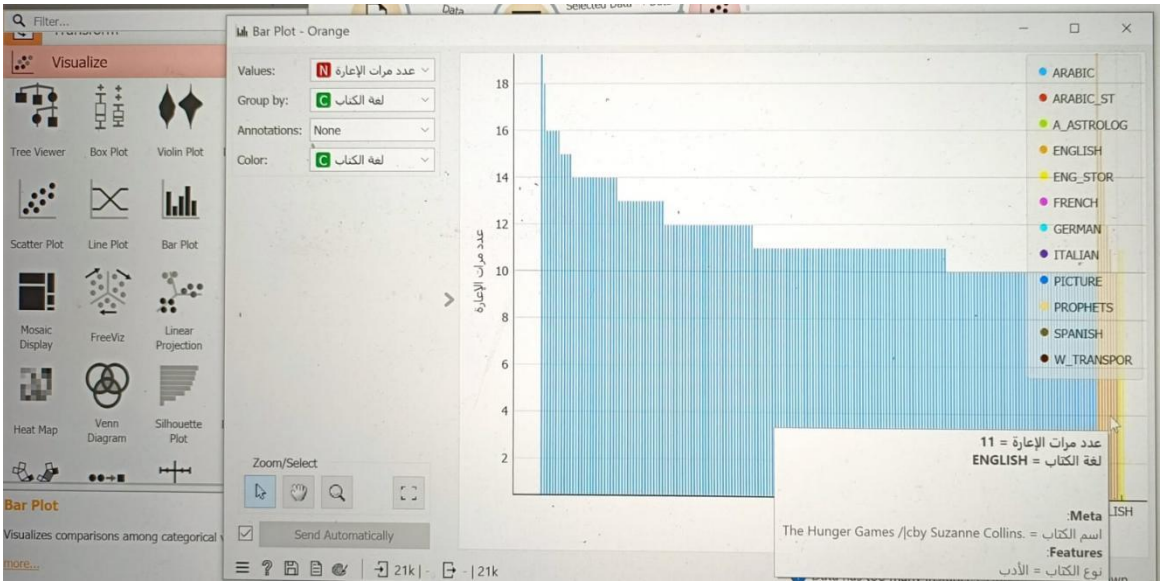
- **الخطوة الثالثة:** هي (تكامل البيانات):

- يمكننا اختيار البيانات المتشابهة وذات الصلة ودمجها معا؛ من خلال الضغط على أداة Scatter Plot وبمجرد الضغط على النقطة يمكننا استخلاص البيانات بشكل منظم؛ كما يتضح في الشكل رقم(7):



شكل رقم (7) تجميع البيانات ذات العلاقة ببعضها البعض.

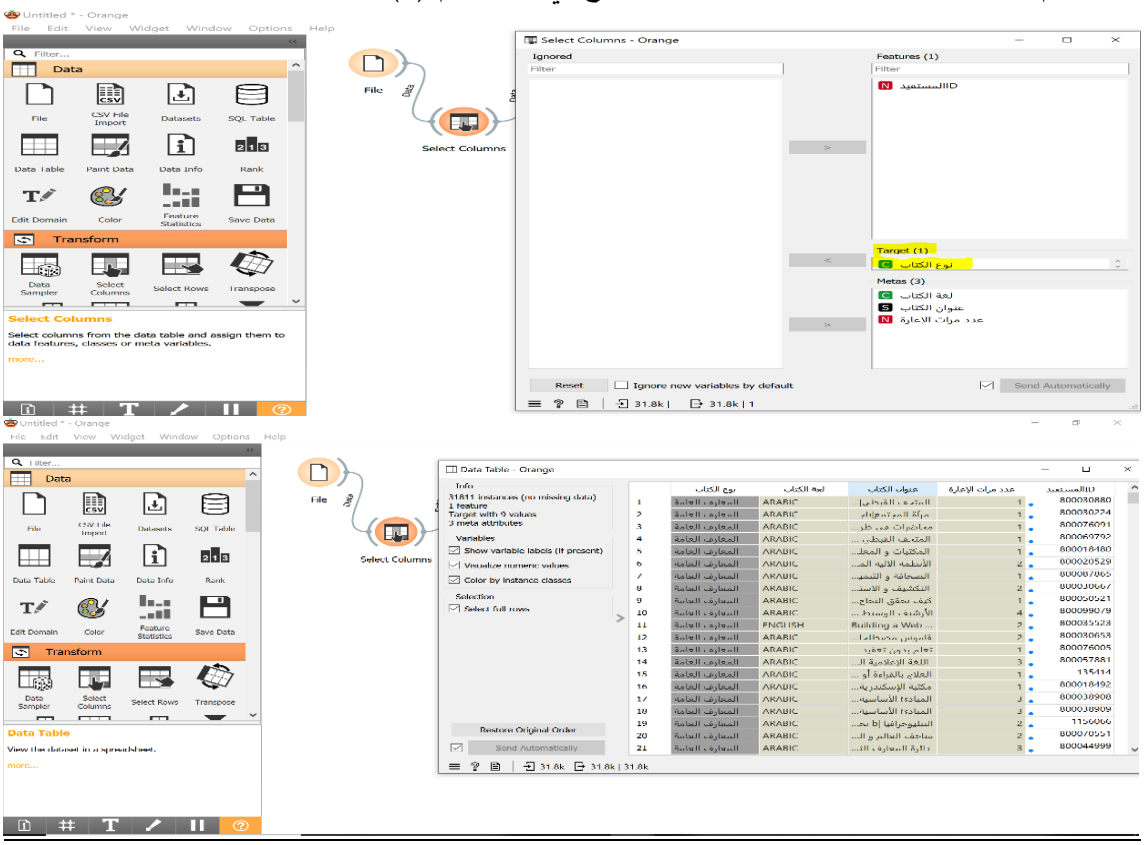
- وفي حالة تنوع البيانات المتشابهة، يمكننا عرضها بأكثر من طريقة من خلال أداة أخرى وهي الـ Barplot؛ كما يتضح في الشكل رقم(8):



شكل رقم(8) عرض البيانات المتشابهة بأداة Bar Plot

- **<u>الخطوة الرابعة</u>**: هي (اختيار البيانات):

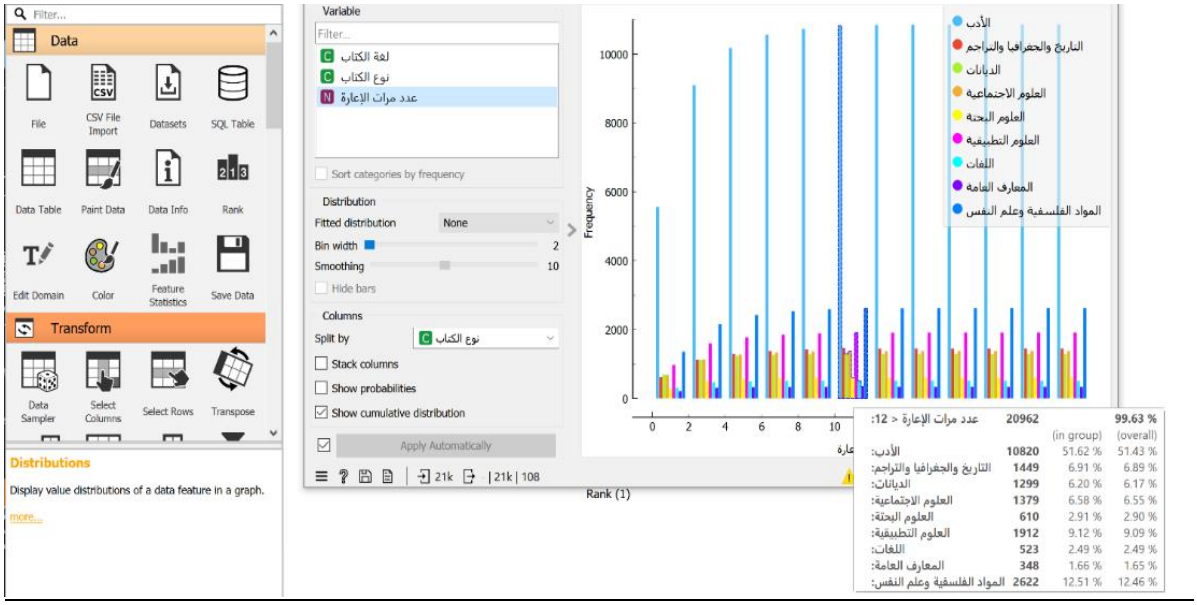
- يتم تحديد واسترجاع البيانات الملائمة أو تحديد أي من هذه البيانات هي Target من مجموعة البيانات التي تم عرضها؛ فيتم اختيار Select Columns؛ كما هو موضح في الشكل رقم (9):

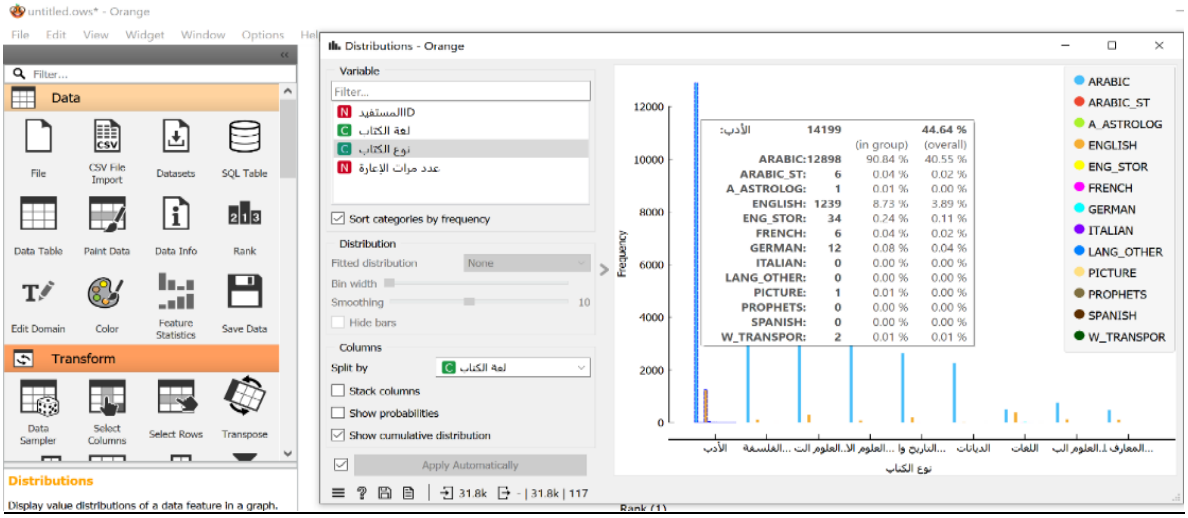


شكل رقم(9) تحديد واسترجاع نوع البيان الـ Target من مجموعة البيانات.

- **<u>الخطوة الخامسة</u>**: هي (تحويل البيانات):

- ويوضح الشكل رقم (10) طريقة تحليل البيانات لفهم خصائصها وتوزيعها لإجراء عمليات البحث والاسترجاع عليها.

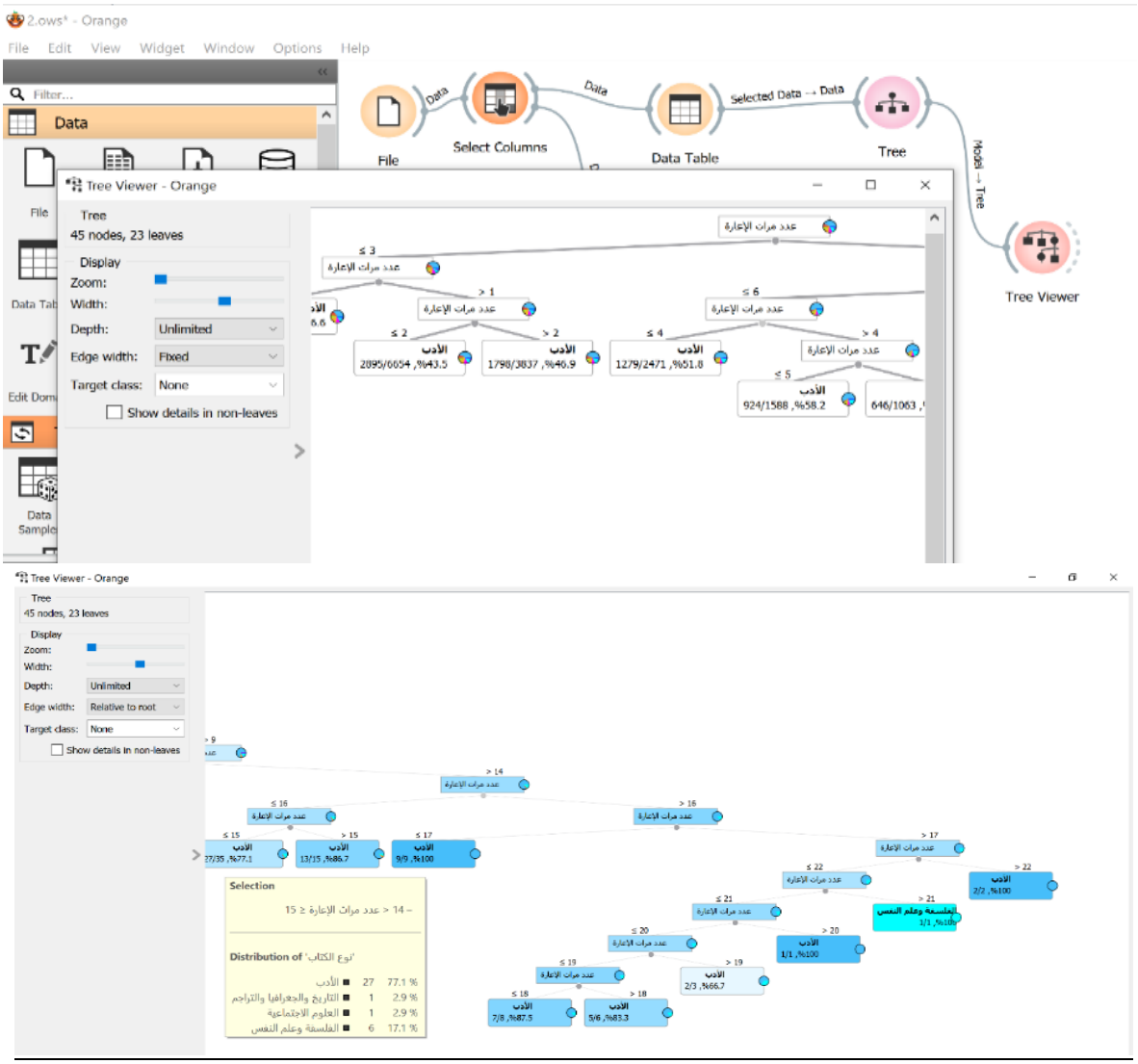




شكل رقم(10) فهم سلوك وخصائص البيانات.

- **<u>الخطوة السادسة</u>**: (التنقيب عن البيانات):

- في هذه الخطوة نقوم ببناء خوارزميات لتمثيل البيانات واستخراج المعرفة؛ وتم في هذه الخطوة استخدام خوارزمية شجرة القرارات كأحد خوارزميات التنقيب عن البيانات كما سبق وذكرنا؛ ويوضحها الشكل رقم(11):



شكل رقم (11) التنقيب عن البيانات باستخدام خوارزمية شجرة القرارات.

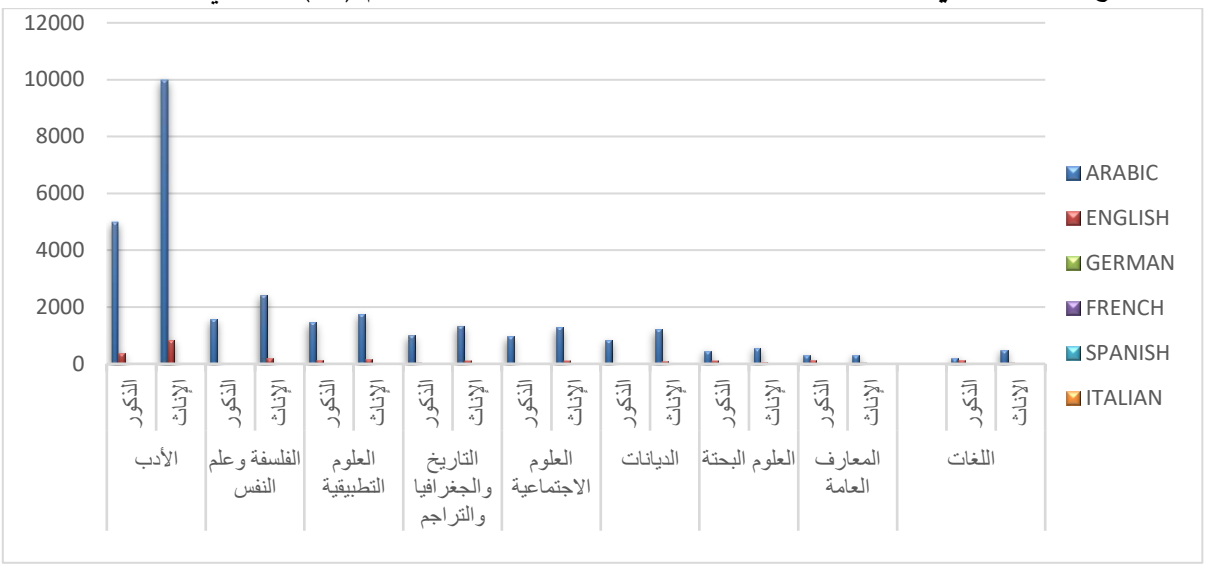
- **وهنا يأتي الهدف الرئيسي من التنقيب عن البيانات وهو** تحويل النتائج المُستخرجة إلى معلوماتٍ واضحة يمكن تطبيقها، والإفادة منها في فهم وتفسير نتائج استخراج البيانات.

رابعا: نتائج وتوصيات الدراسة:

توصلت الدراسة لمجموعة من النتائج والتوصيات، ويمكن حصر أبرز هذه النتائج في النقاط التالية:

- يُعد برنامج Orange3 من أكثر الأدوات شيوعًا في مجال التنقيب عن البيانات وتحليليها؛ لأنه يوفر مجموعة من أدوات التحليل المُتقدمة التي تدعم عمليات البحث في البيانات المهيكلة وغير المهيكلة، كما يقوم بعرض وتنظيم مجموعات كبيرة من البيانات الخاصة أو العامة بسهولة، واستخراج الكيانات والعلاقات منها للوصول إلى معلومات واضحة يمكن الإفادة منها.
- تمثلت نتائج التنقيب في ملفات بيانات الإعارة متمثلة في تقسيم المستفيدين إلى إناث وذكور وفقا للمراحل العمرية المختلفة؛ كما هو موضح في الآتي:

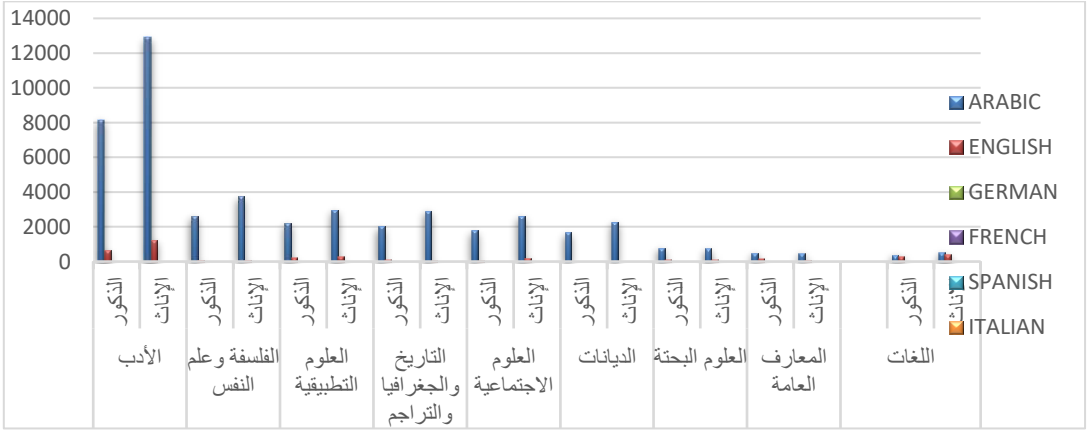
➢ **نتائج المرحلة العمرية من 16 سنة وحتى سن 21 سنة:** يوضحها الشكل رقم (12)؛ كما يلي:



شكل رقم (12) سمات المرحلة العمرية من سن 16- 21

- إجمالي تسجيلات الإعارة في تلك المرحلة العمرية للإناث هو21042، وللذكور 12725 تسجيلة.
- بالنسبة للإناث؛ كانت مصادر المعلومات باللغة العربية أكثر إعارةً من اللغات الأخرى بنسبة 91%، ويليها اللغة الإنجليزية بنسبة 9%، أما بالنسبة للذكور؛ فكانت اللغة العربية أيضا أكثر إعارة من اللغات الأخرى بنسبة 92%، ويليها مصادر المعلومات باللغة الإنجليزية بنسبة 8%.
- كما كانت أكثر المجالات المعرفية إعارة للإناث هو مجال الأدب بنسبة 52%، يليه مجال الفلسفة وعلم النفس بنسبة 12%، أما الذكور؛ فشكل مجال الأدب نسبة 43%، يليه الفلسفة وعلم النفس والعلوم التطبيقية بنسبة 13%.
- أكثر مصادر المعلومات إعارة بالنسبة للإناث؛ هو كتاب " MySQL Web "Beginning PHP5, Apache, and development بعدد مرات (23) إعارة، وكان هناك أكثر من 10756 مصدر معلومات تم إعارتها لمرة واحدة فقط.
- أما بالنسبة للذكور؛ فكان أكثر مصادر المعلومات إعارة هو كتاب " كفاحي" للمؤلف أدولف هتلر بعدد مرات (13) مرة، ووجد أكثر من 9432 مصدر للمعلومات تم إعارتها لمرة واحدة فقط.

➢ **نتائج المرحلة العمرية من سن 21 سنة وحتى سن 26 سنة:** يوضحها الشكل رقم (13)؛ كما يلي:

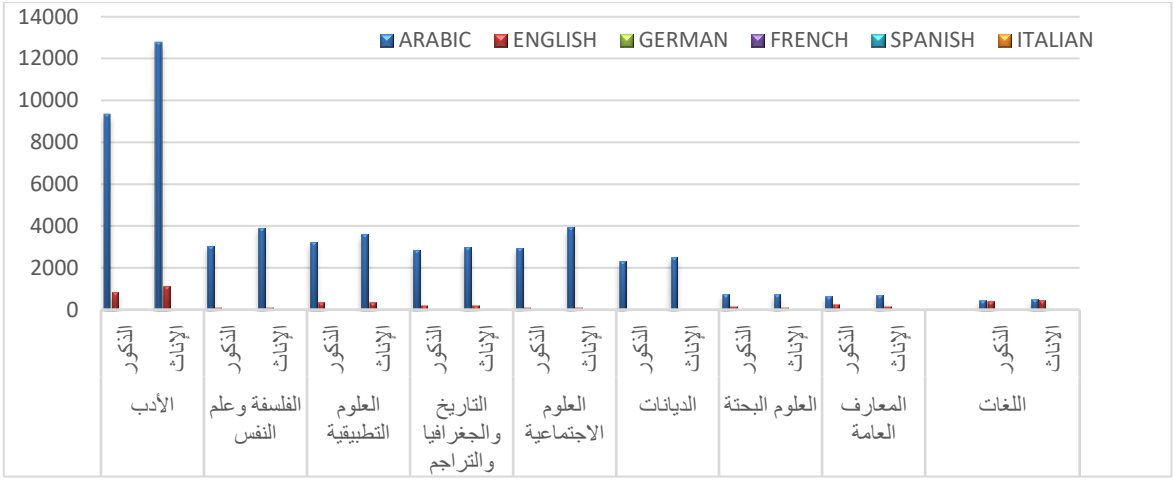


شكل رقم (13) سمات المرحلة العمرية من سن 21-26

- إجمالي تسجيلات الإعارة في تلك المرحلة العمرية للإناث هو 31812، وللذكور21922 تسجيلة.
- بالنسبة للإناث؛ كانت أكثر المجالات المعرفية إعارة هو مجال الأدب بنسبة 53%، يليه مجال الفلسفة وعلم النفس بنسبة 15%، وشكلت المعارف العامة، والعلوم البحتة، واللغات أقل المجالات الموضوعية إعارة.

- أما الذكور؛ فكان أكثر المجالات المعرفية إعارة هو مجال الأدب بنسبة 42%، يليه مجال الفلسفة وعلم النفس والعلوم التطبيقية بنسبة 14%.
- شكلت مصادر المعلومات باللغة العربية للإناث النسبة الأكثر إعارة من اللغات الأخرى بنسبة 93%، يليها اللغة الإنجليزية بنسبة7%، أما بالنسبة للذكور؛ فشكلت نسبة مصادر المعلومات باللغةِ العربية أيضا الأكثر إعارة من اللغات الأخرى بنسبة 92%، يليها اللغة الإنجليزية بنسبة 8%.
- كانت رواية " ذاكرة الجسد" لأحلام مستغانمي من أكثر مصادر المعلومات إعارةً في هذه المرحلة للإناث بعدد مرات (25)، يليها رواية "قد تكون قصة حب" لمارتن باج بعدد (23) مرة، وكان هناك أكثر من 13907 مصادر معلومات تم إعارتها لمرة واحدة فقط، أما الذكور فكان كتاب "النظرية النسبية الخاصة" أكثر مصادر المعلومات إعارة بعدد مرات (17)، بينما كان هناك حوالي 12767 مصادر معلومات تم إعارتها لمرة واحدة فقط.

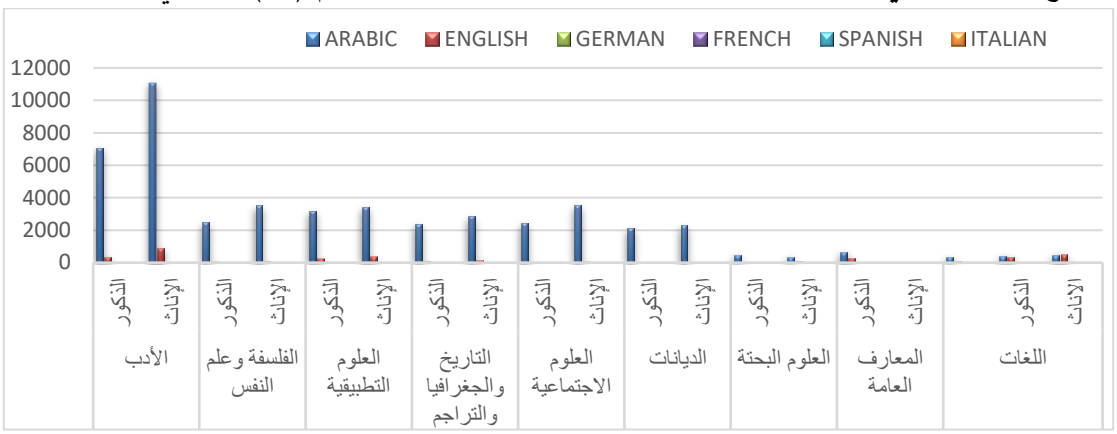
➢ **<u>نتائج المرحلة العمرية من 26 سنة وحتى سن 31 سنة:</u>** يوضحها الشكل رقم (14) كما يلي:



شكل رقم (14) سمات المرحلة العمرية من سن 26-31

- شكلت تلك المرحلة أكثر المراحل العمرية إقبالا على قسم الإعارة بإجمالي عدد تسجيلات 34305 للإناث وعدد 27978 تسجيلة للذكور.
- بالنسبة للإناث والذكور أيضا؛ كانت مصادر المعلومات باللغة العربية أكثر إعارةً من اللغات الأخرى بنسبة 92%، ويليها اللغة الإنجليزية بنسبة 8%.
- ويوضح الشكل رقم (15)؛ أن أكثر المجالات المعرفية إعارةً بالنسبة للإناث؛ هو مجال الأدب بنسبة 46%، يليه مجال الفلسفة وعلم النفس بنسبة 14%، أما بالنسبة للذكور؛ فكان أكثر المجالات المعرفية إعارة هو مجال الأدب بنسبة 37%، يليه العلوم التطبيقية بنسبة 14%.
- كان كتاب "بوصلة الشخصية" للمؤلفين دايان تيرنر، وثلما جريكو أكثر الكتب إعارة بالنسبة للإناث بعدد مرات (24)؛ يليه قصة "النداهة" ليوسف إدريس بعدد مرات إعارة (22) مرة، وكان هناك حوالي 14905 مصادر معلومات قد تم إعارتها لمرة واحدة فقط، أما بالنسبة للذكور؛ فكان كتاب " قواعد النجاح والسعادة في حياتك العملية" للمؤلفة سوزان موري أكثر إعارة بعدد مرات (21)، يليه كتاب " صناعة الشباب" لمؤلفه محمد سعيد حوي بعدد (17) مرة، وكان هناك عدد 14321 مصدر معلومات قد تم إعارتها لمرة واحدة فقط.

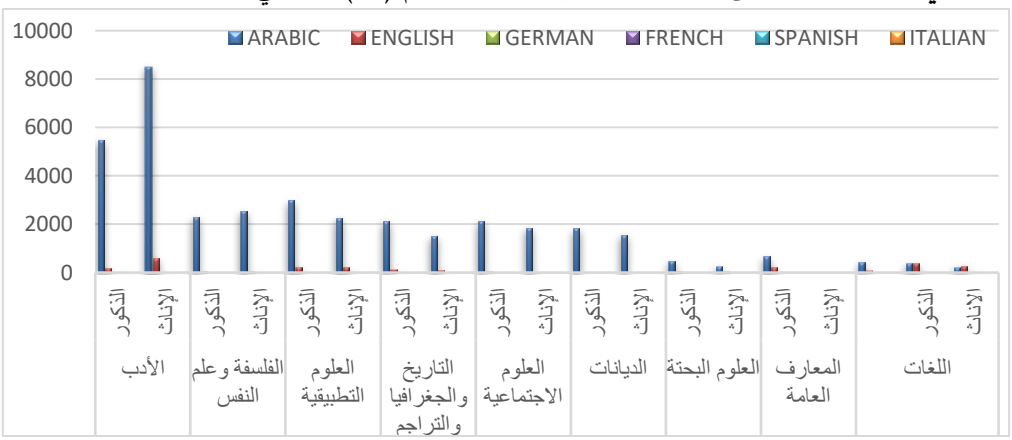
➢ **<u>نتائج المرحلة العمرية من 31 سنة وحتى سن 36 سنة:</u>** يوضحها الشكل رقم (15)؛ كما يلي:



شكل رقم (15) سمات المرحلة العمرية من سن 31-36

- إجمالي عدد تسجيلات الإعارة للذكور في تلك الفئة العمرية 25134، بينما وصل عدد تسجيلات الإعارة للإناث30563 تسجيلة.
- كانت أكثر المجالات المعرفية إعارة بالنسبة للإناث؛ هو مجال الأدب بنسبة 42%، يليه مجال الفلسفة وعلم النفس بنسبة 15%، أما بالنسبة للذكور؛ فكان أكثر المجالات المعرفية إعارة هو مجال الأدب بنسبة 36%، يليه العلوم التطبيقية بنسبة 14%.
- بالنسبة للإناث والذكور أيضا؛ كانت مصادر المعلومات باللغة العربية أكثر إعارةً من المصادر باللغات الأخرى بنسبة 93%، ويليها اللغة الإنجليزية بنسبة 7%.
- كان كتاب "أسعد امرأة في العالم" للمؤلف عائض عبد الله القرني أكثر الكتب إعارة بالنسبة للإناث بعدد مرات (23)؛ يليه كتاب " تبدأ الحياة عندما تبدؤها أنت" للمؤلفة ماري آن رادماتشر بعدد مرات إعارة (19) مرة، وكان هناك 14397 مصدر معلومات تم إعارتها لمرة واحدة فقط، أما بالنسبة للذكور؛ فكان كتاب "Aircraft electricity & electronics" للمؤلف Thomas K. Eismin أكثر إعارة بعدد مرات (31) مرة، يليه كتاب " يوميات إسكندرية" لأمل الجيار بعدد (21) مرة، وكان هناك حوالي 13848 مصدر معلومات قد تمت إعارتها لمرة واحدة فقط.

➢ **<u>المرحلة العمرية من 36 سنة وحتى سن 41 سنة:</u>** يوضحها الشكل رقم (16)؛ كما يلي:

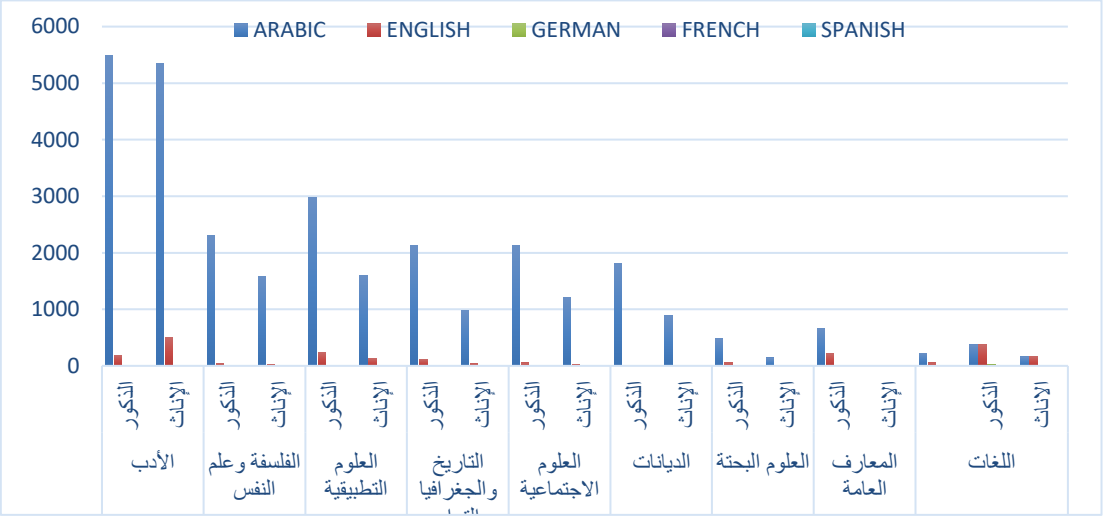


شكل رقم (16) سمات المرحلة العمرية من سن 36-41

- إجمالي عدد تسجيلات الإعارة للذكور في تلك الفئة العمرية 22741، بينما وصل عدد تسجيلات الإعارة للإناث23654 تسجيلة.

- كانت أكثر المجالات المعرفية إعارة بالنسبة للإناث؛ هو مجال الأدب بنسبة 43%، يليه مجال الفلسفة وعلم النفس بنسبة 16%، أما بالنسبة للذكور؛ فكان أكثر المجالات المعرفية إعارة هو مجال الأدب بنسبة 30%، يليه العلوم التطبيقية بنسبة 17%.
- بالنسبة للإناث والذكور أيضا؛ كانت مصادر المعلومات باللغة العربية أكثر إعارةً من الكتب باللغات الأخرى بنسبة 93%، ويليها اللغة الإنجليزية بنسبة 7%.
- كان كتاب "أماكن في القلب" للمؤلف عبد الوهاب مطاوع أكثر الكتب إعارة بالنسبة للإناث بعدد مرات (14) مرة؛ يليه كتاب " تعالوا نربي: روشتة في تربية أطفالنا" للمؤلفة أمنية عادل عبد الرازق بعدد مرات إعارة (11) مرة، وكان هناك 13547 مصدر معلومات تم إعارتها لمرة واحدة فقط، أما بالنسبة للذكور؛ فكان كتاب " فن تعليم قيادة السيارات" للمؤلفة مارجريت ستاسي أكثر إعارة بعدد مرات (30) مرة، يليه كتاب "Android phones for dummies" للمؤلف Dan Gookin بعدد (21) مرة، وكان هناك أكثر من 13848 مصدر معلومات تم إعارتها لمرة واحدة فقط.

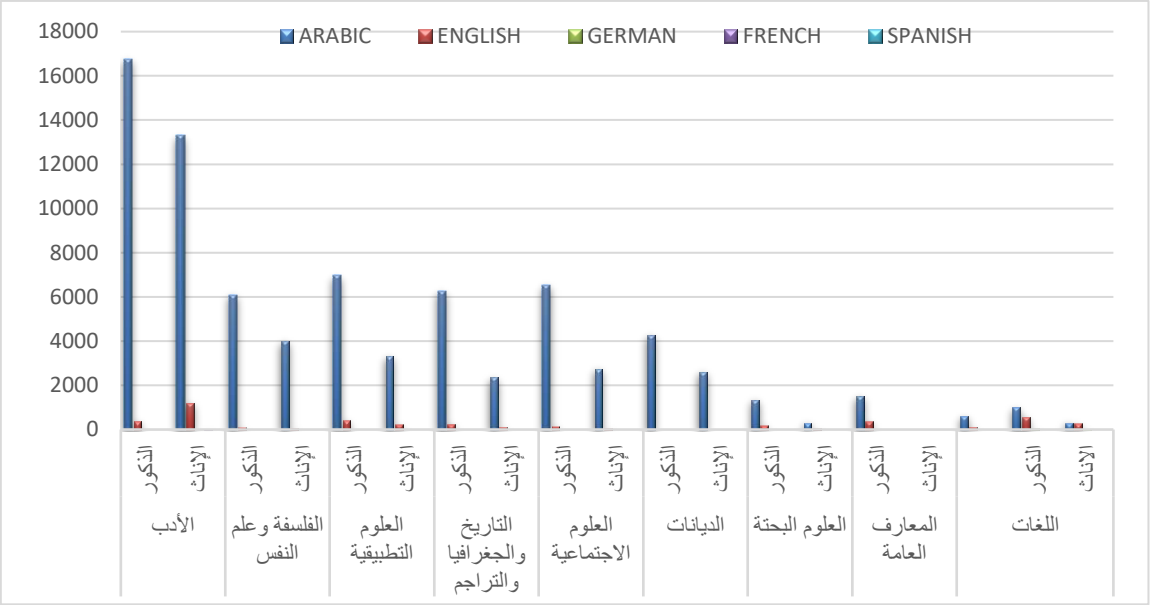
➢ **<u>المرحلة العمرية من 41 سنة وحتى سن 46 سنة:</u>** يوضحها الشكل رقم (17)؛ كما يلي:



شكل رقم (17) سمات المرحلة العمرية من سن 41-46

- إجمالي تسجيلات الإعارة في تلك المرحلة العمرية للإناث هو 20572، وللذكور 19712 تسجيلة.
- كانت أكثر المجالات المعرفية إعارة بالنسبة للإناث؛ هو مجال الأدب بنسبة 46%، يليه مجال الفلسفة وعلم النفس بنسبة 14%، أما بالنسبة للذكور؛ فكان أكثر المجالات المعرفية إعارة هو مجال الأدب بنسبة 25%، يليه العلوم التطبيقية بنسبة 18% وتراوحت باقي النسب الأخرى على مجالات المعرفة المختلفة.
- بالنسبة للإناث؛ كانت مصادر المعلومات باللغة العربية أكثر إعارةً من الكتب باللغات الأخرى بنسبة 94%، ويليها اللغة الإنجليزية بنسبة 6%، أما بالنسبة للذكور؛ فشكلت نسبة مصادر المعلومات باللغة العربية أيضا الأكثر إعارة من اللغات الأخرى بنسبة 93%، يليها اللغة الإنجليزية بنسبة7%.
- شكلت الروايات النسبة الأكبر للإعارة بالنسبة للإناث؛ فنجد أن رواية "أنا شهيرة: الحكاية الأولى" للمؤلفة نور عبد الحميد قد تمت إعارتها بعدد (10) مرات، يليها رواية "باب الأسرار" للمؤلف أحمد أوميت بعدد (10) مرات، وكان هناك عدد 12743 مصدر معلومات قد تم إعارتها لمرة واحدة فقط، أما بالنسبة للذكور؛ فكان كتاب " أساسيات الفيزياء" لبوش أكثر إعارة بعدد مرات (29) مرة، يليه كتاب " تكنولوجيا الاتصال ومجتمع المعلوماتية" لحنان يوسف بواقع (24) مرة إعارة، وكان هناك عدد 12631 مصدر معلومات قد تم إعارتها لمرة واحدة فقط.

➢ **المرحلة العمرية من 46 سنة وحتى سن 60 سنة:** يوضحها الشكل رقم (17)؛ كما يلي:



شكل رقم (18) سمات المرحلة العمرية من سن 46-60

- بالنسبة للإناث؛ كانت أكثر المجالات المعرفية إعارة مجال الأدب بنسبة 30%، وتساوت كل من العلوم التطبيقية والعلوم الاجتماعية بنسبة14%، أما بالنسبة للذكور؛ فكان أكثر المجالات المعرفية إعارةً هو مجال الأدب بنسبة 30%، يليها العلوم التطبيقية والعلوم الاجتماعية بنسبة 14%، وتراوحت باقي النسب الأخرى على مجالات المعرفة المختلفة.
- بالنسبة للإناث؛ كانت مصادر المعلومات باللغة العربية أكثر إعارةً من الكتب باللغات الأخرى بنسبة 94%، ويليها اللغة الإنجليزية بنسبة 6%، أما للذكور؛ فشكلت نسبة مصادر المعلومات باللغة العربية أيضا الأكثر إعارة من اللغات الأخرى بنسبة 95%، يليها اللغة الإنجليزية بنسبة5%.
- كان كتاب " تاريخ أكثر إيجازا للزمن" للمؤلفين ستفن هوكينج، وليونارد ملوندينوف أكثر إعارة بالنسبة للإناث بعدد مرات (27) مرة؛ يليه كتاب " digital classroom" للمؤلف Jennifer Smith بعدد مرات إعارة (15) مرة، أما بالنسبة للذكور؛ فكان كتاب " The human genom" للمؤلفين Jeremy Cherfas, John Griblin أكثر إعارة بعدد مرات (29) مرة إعارة، يليه كتاب" الصيدلية الخضراء: دليل العلاج بالأطعمة" للمؤلف جيمس ديوك بعدد (19) مرة إعارة.
- كما يمكننا الخروج بمؤشرات عامة للميول القرائية والبحثية للمراحل العمرية ككل؛ تمثلت في الآتي:
  - كانت الإناث أكثر ميلا لاستعارة مصادر المعلومات في مجال الفلسفة وعلم النفس، أما الذكور فكانت أكثر ميلا لاستعارة مصادر المعلومات في مجال العلوم التطبيقية.
  - اتجه كل من الإناث والذكور نحو استعارة مصادر المعلومات باللغة العربية أكثر من اللغات الأخرى، ثم اللغة الإنجليزية.
  - شكلت المرحلة العمرية من سن 26 سنة وحتى سن 31 سنة للذكور والإناث أكثر المراحل العمرية إعارةً للروايات والقصص بالمكتبة.

ومن ثم يمكننا تقديم مجموعة الاستنتاجات حول الإفادة من تطبيق آليات التنقيب عن البيانات في تحسين مستوى خدمة الإعارة بالمكتبات ومؤسسات المعلومات، بناءً على التعاملات التي تحدث بين المستفيد ومصادر المعلومات؛ والتي من خلالها يمكننا تحقيق ما يلي:

- القدرة على معرفة الميول القرائية والاتجاهات البحثية لفئة معينة من المستفيدين بمختلف أعمارهم وثقافاتهم.
- القدرة على تحديد العلاقة بين المستفيدين وأنواع المواد المستعارة.
- القدرة على تحديد قائمة المواد الأكثر استعارة والتي يتم الطلب عليها بشكل متكرر.
- القدرة على تحديد قائمة المواد التي لم يتم الطلب عليها؛ وبالتالي يتم استبعادها.
- تحديد أكثر مصادر المعلومات والموضوعات الأكثر استخداما على مستوى مجال معرفي معين.
- تقديم المقترحات لمديري ومتخذي القرار بالمكتبات ومؤسسات المعلومات حول إعادة التصور لترتيب مقتنيات المكتبة، وبناء التوقعات المستقبلية نحو تداول مقتنيات المكتبة في فترات زمنية معينة، وإعادة النظر في إستراتيجيات تزويد مقتنيات جديدة، وكذلك وضع سياسات جديدة لمقتنيات المكتبة الموجودة بالفعل، إضافًة إلى ترشيد أوجه صرف الميزانية.

ومن هذا المنطلق توصي هذه الدراسة بالآتي:

- توجيه المكتبات ومؤسسات المعلومات نحو الاستفادة من تطبيق آليات التنقيب عن البيانات وخوارزمياتها في فهم الاحتياجات القرائية والبحثية للمستفيدين منها؛ وذلك لتحسين جودة الخدمات المقدمة وتقديم الخدمات الفعلية التي يحتاجونها.
- عقد ورش العمل لرفع وعي العاملين بالمكتبات ومؤسسات المعلومات بأدوات التنقيب عن البيانات، وكيفية تطبيقها بمكتباتهم لتحقيق أقصى استفادة ممكنة.
- حث الباحثين على إجراء المزيد من الدراسات التطبيقية حول استخدام آليات التنقيب عن البيانات في تقديم خدمات المعلومات بالمكتبات ومؤسسات المعلومات؛ وذلك من خلال تحليل استفسارات المستفيدين عند تقديم الخدمة المرجعية والرد على الاستفسارات.
- توجيه أقسام المكتبات بالجامعات المصرية والعربية لتضمين مقرر التنقيب عن البيانات ضمن مقررات برامج علوم المكتبات.
- توجيه المكتبات ومؤسسات المعلومات نحو الاستفادة من تجارب المكتبات العالمية التي طبقت تقنيات التنقيب عن البيانات في فهم احتياجات المستفيدين، وتقديم خدماتها بكفاءة وجودة عالية.

الخاتمة:

تناولت هذه الدراسة توضيح آلية تطبيق أساليب التنقيب عن البيانات في تحليل بيانات ملفات الإعارة للكبار في مكتبة مصر العامة، وذلك من خلال عرض للمفاهيم الرئيسية حول التنقيب عن البيانات وتطبيقاته وقضاياه الأساسية ومهامه، وتوضيح الأساليب والطرق المستخدمة في التنقيب عن البيانات، وديناميكية عمل كل منها، ثم تطرقت إلى عرض وتحليل برنامج Orange3 وتوضيح سير العمل به، وخطوات تطبيقه في استخراج المعرفة من تلك الملفات، وأخيرًا كشفت الدراسة عن طرق الاستفادة من تلك المعرفة في الكشف عن الاحتياجات القرائية والبحثية للمستفيدين من المكتبة؛ لتوجيه المقترحات لمديري المكتبة ومتخذي القرار نحو اقتناء مصادر المعلومات التي تلبي تلك الاحتياجات، وتحسين كفاءة وجودة الخدمات المقدمة لهم.

وخلُصت الدراسة أنَّه يجب توجيه المكتبات ومؤسسات المعلومات نحو الإفادة من تقنيات التنقيب عن البيانات في تقديم خدماتها بكفاءة وجودة عالية؛ وذلك لما توفره من معلومات حول المستفيدين منها وتحديد الاحتياجات القرائية لهم، مَما يُسهم في تقديم رؤية جديدة حول ما تقتنيه من مصادر معلومات، وما تُقدمه من خدمات تلبي احتياجات تلك المستفيدين، كما يمكن الإفادة مِن أدوات التنقيب عن البيانات في بناء قاعدة بيانات تُساعد العاملين في المكتبات ومؤسسات المعلومات ومتخذي القرار من استنباط رؤى جديدة لتنمية مجموعات المكتبة والخدمات المستقبلية التي يجب أن تقدمها.

المراجع

المراجع العربية:

- أحمد، أحمد عبد المطلب& عبد الله، محمد حسن، (مارس،2018)، تنقيب البيانات التعليمية باستخدام خوارزمية APRIORI لتقييم أداء وتحسين الإنتاج الأكاديمي، مجلة النيل الأبيض للدراسات والبحوث، 1،11-37.
http://search.mandumah.com/Record/962864
- بابكر، خولة علي & علي، الأمير خير الله الأمين،(2014) ، تقنيات البحث والتنقيب في مستودعات بيانات الويب بالتطبيق على البحث عن الوظائف (رسالة ماجستير جامعة أم درمان الاسلامية).
http://search.mandumah.com/Record/651399
- حسن، فارس جمعة & زين العابدين، عمار عبد اللطيف، (ديسمبر،2023)، مجالات التنقيب عن البيانات في المؤسسات التعليمية: المكتبات الجامعية نموذجا، آداب الرافدين، 53(95)، 304-319.
http://search.mandumah.com/Record/1480191
- خوجلي، ريان الفاضل صديق& عبد الله، نفيسة حسن محمد، (2018) ، استخدام تقنيات تنقيب البيانات لدراسة ظاهرة تسرب الطلاب في مدارس تعليم الأساس بالتطبيق على منطقتي جنوب الجزيرة والحص حيصا ولاية الجزيرة (رسالة ماجستير، جامعة النيلين).http://search.mandumah.com/Record/1140687
- درويش، وسام محمود أحمد، (يونيو،2009)، نحو رؤية جديدة لإدارة المكتبات باستخدام تقنية التنقيب عن البيانات(Data Mining) . Cybrarians Journal، 19،19-39.
http://search.mandumah.com/Record/507947
- رجب، نهى حسن عثمان& فتوح، سيف الدين& أحمد، عوض حاج علي، (2018) ، تحسين خوارزميات تعدين التنبؤ في الأداء الأكاديمي للطلاب: حالة تطبيقية جامعة بحري (رسالة دكتوراه، جامعة النيلين).
http://search.mandumah.com/Record/1033740.
- زكي، فاطمة يحيى، (سبتمبر،2020)، القضايا الأخلاقية والقانونية لاستخدام بيانات المستفيدين للتنقيب عن البيانات في المكتبات المصرية: دراسة تحليلية.Cybrarians Journal، 59، 1-52.
http://search.mandumah.com/Record/1093383
- السعيدي، هيلة عبد الله& الميلبي، أروى بنت نصار، (نوفمبر،2021)، منهج بناء خرائط المعرفة المستند على تقنيات التنقيب في البيانات لاتخاذ القرار بالمؤسسات: دراسة وصفية تطويرية، اعلم، 30، 289-310.
http://search.mandumah.com/Record/1338760
- سيد، أحمد فايز أحمد، (2016)، أدوات التنقيب عن البيانات مفتوحة المصدر: دراسة تحليلية تقييمية، مجلة جامعة طيبة للآداب والعلوم الإنسانية، 5(10)، 792 -865.http://search.mandumah.com/Record/773730
- ـــــــــــــــــ، (يونيو،2013)، التنقيب عن بيانات مؤسسات العمل التطوعي على الويب: دراسة تحليلية، مجلة كلية الآداب، ع27، 355-442. http://search.mandumah.com/Record/1216093
- علوان، سها السيد منصور السيد، (ديسمبر2020)، تنقيب البيانات في مجال المكتبات والمعلومات: مراجعة علمية، المجلة الدولية لعلوم المكتبات والمعلومات، 7(4)، 373 -396.
http://search.mandumah.com/Record/1106688
- لبيب، نيفين مكرم& إبراهيم، عادل، (2003)، اكتشاف المعرفة من خلال التنقيب على البيانات والنصوص والتطبيق على التطبيقات الطبية، في المؤتمر العلمي العاشر لنظم المعلومات وتكنولوجيا الحاسبات: إدارة المعرفة في حقبة العولمة (ص 104)، الجمعية المصرية لنظم المعلومات وتكنولوجيا الحاسبات ومركز البحوث والمعلومات بأكاديمية السادات للعلوم الإدارية. http://search.mandumah.com/Record/503840
- المبارك، البدوي سعد البدوي& إبراهيم، أشرف عثمان،(2017) ، استخدام تقنيات تنقيب البيانات لاستكشاف أنماط مؤثرات التحصيل الأكاديمي لطلاب المرحلة الثانوية: دراسة حالة شهادة التعليم الثانوي الصف الثالث علمي من العام 2007م حتى العام 2016م (رسالة ماجستير، جامعة النيلين).
http://search.mandumah.com/Record/915822

- محمد، محمد إبراهيم حسن، (يوليو،2011)، بناء أنطولوجيات التنقيب عن البيانات: تحليل لآليات التنفيذ في بيئة مفتوحة المصدر(1) ، مجلة المكتبات والمعلومات العربية، 31(3)، 31-86. http://search.mandumah.com/Record/111783
- موسى، صباح محمد & عبد الحسن، صادق حسين، (2015)، أثر تقنية تنقيب البيانات في تطوير عملية صنع القرارات: دراسة استطلاعية في وزارة التعليم العالي والبحث العلمي العراقية، مجلة الإدارة والاقتصاد، 38(103)، 69-84. http://search.mandumah.com/Record/803195
- النشرتي، مؤمن سيد، (مارس،2014)، نحو التكامل المعرفي من واقع توظيف الأنطولوجيات في أطار التنقيب عن البيانات: دراسة تحليلية،Cybrarians Journal ، 34 ، 162-198. http://search.mandumah.com/Record/510946

المراجع الأجنبية:

- Ansari, N. & Vakilimofrad, H.& Mansoorizadeh, M. & Amiri, M.R. (2021, October). Using data mining techniques to predict user's behavior and create recommender systems in the libraries and information centers. Global Knowledge Memory and Communication, 70 (6-7), 538-557. https://0810b0ns3-1106-y-https-doi-org.mplbci.ekb.eg/10.1108/GKMC-04-2020-0058.
- Bin, C. (2013). Study on Data Mining in Digital Libraries. In Yang, Y., Ma, M., Liu, B. (eds) Information Computing and Applications. ICICA 2013. Communications in Computer and Information Science, 392. https://doi.org/10.1007/978-3-642-53703-5_30.
- Cambridge University Press(1999). Cambridge Dictionaries Online. https://dictionary.cambridge.org/dictionary/english/artificial-intelligence.
- Chen, S.& Xue, Y. & Cui, X. (2024, February). Information literacy of college students from library education in smart classrooms: based on big data exploring data mining patterns using Apriori algorithm. Soft Computer 28, 3571–3589. https://doi.org/10.1007/s00500-023-09621-8.
- Geoff Mulgan.(2018, September). Artificial intelligence and collective intelligence: the emergence of a new field. AI & SOCIETY, 33(4). https://doi.org/10.1007/s00146-018-0861-5.
- Guirong Lin, Anying Miao. (2022, July). Construction of Library Personalized Intelligent Service System Based on Data Mining. In IPEC '22: Proceedings of the 3rd Asia-Pacific Conference on Image Processing, Electronics and Computers. 857 – 862. https://doi.org/10.1145/3544109.3544368.
- Han, C.& Yu, W.& Li, X.& Lin, H.& Zhao, H. (2022). A New Fast Algorithm for Library Circulation Data Mining Based on FUP. Scientific Programming. https://doi.org/10.1155/2022/1683099.
- Han, Jiawei &Kamber, Micheline& Pei, Jian. (2012). Data Mining: concepts and techniques. Elsevier. https://myweb.sabanciuniv.edu/rdehkharghani/files/2016/02/The-Morgan-Kaufmann-Series-in-Data-Management-Systems-Jiawei-Han-Micheline-Kamber-Jian-Pei-Data-Mining.-Concepts-and-Techniques-3rd-Edition-Morgan-Kaufmann-2011.pdf.
- Hossain, M. & Rahman, R. M. (2015). Application of Data Mining Techniques on Library Circulation Data for Library Material Acquisition and Budget Allocation. In M. Tavana & K. Puranam (Eds.), Handbook of Research on Organizational Transformations through Big Data Analytics, 334-354, https://doi.org/10.4018/978-1-4666-7272-7.ch020.
- Hu, J.& Tang, J.& Zhang, B. (2019). Research on Readers' Behavior of College Libraries Based on Data Mining Technology. Journal of Physics: Conference Series, 1345(5). https://doi.org/10.1088/1742-6596/1345/5/052010.

- Huancheng, L.& Tingting, W. & Rocha, Á. (2019, September). An Analysis of Research Trends on Data Mining in Chinese Academic Libraries. J Grid Computing, 17, 591–601. https://doi.org/10.1007/s10723-018-9461-3
- Izquierdo, H. (2022). Artificial Intelligence and Text and Data Mining: Future Rules for Libraries. In J. Coates, V. Owen & S. Reilly (Ed.), Navigating Copyright for Libraries: Purpose and Scope ,497-540. https://doi.org/10.1515/9783110732009-022
- J. Song. (2022). Functions and Ways of Computer Information Technology in Optimizing Library Operation and Management Considering Massive Data Mining. In 6th International Conference on Computing Methodologies and Communication (ICCMC), Erode, India,717-720. doi:https://081040nd7-1106-y-https-doi-org.mplbci.ekb.eg/10.1109/ICCMC53470.2022.9754122.
- J, Yu. (2021). Construction of Smart Library Based on the Big Data Mining and Knowledge Discovery. Smart City Challenges & Outcomes for Urban Transformation (SCOUT), Bhubaneswar, India, 57-61. doi:https://081040nd7-1106-y-https-doi-org.mplbci.ekb.eg/10.1109/SCOUT54618.2021.00023.
- Katsurai, Marie& Joo, Soohyung. (2021, June). Adoption of data mining methods in the discipline of library and information science. Journal of Library and Information Studies, 19(1), 1 – 17. Available at: https://jlis.lis.ntu.edu.tw/files/journal/j52-1.pdf.
- Kaderye, Golam & Arif, Ahsan& Kundu, Ronjon. (2024, March). Data Mining in Different Fields: A Study. International Journal of Innovative Science and Research Technology,9(3).: https://doi.org/10.38124/ijisrt/IJISRT24MAR1384.
- Khademi, S., Rafieinasab, F., & Radhakrishnan, N. (2023). Data Mining in Academic Libraries: A Systematic Review. International Journal of Information Science and Management (IJISM), 21(3), 255-271. https://doi.org/10.22034/ijism.2023.1977879.0
- Leilei Peng, Ke Chen. (2023). Mining library virtual reference service data by using the software NVivo 12. In 2nd International Conference on Computer Graphics, Artificial Intelligence, and Data Processing (ICCAID 2022). https://doi.org/10.1117/12.2674684
- Li, Zhenxin. (2022). Digital libraries archive information integration based on multidimensional data mining. International Journal of Reasoning-based Intelligent Systems (IJRIS), 14(4). https://doi.org/ 10.1504/IJRIS.2022.126659
- Liu, Y. (2018, February). Data Mining of University Library Management Based on Improved Collaborative Filtering Association Rules Algorithm. Wireless Pers Commun, 102, 3781–3790. https://doi.org/10.1007/s11277-018-5409-y.
- Lone, T. A& Khan, R.A. (2014). Data mining: Competitive tool to digital library. Journal of Library and Information Technology, 34(5), 401– 406. https://doi.org/10.14429/djlit.34.5.6722
- Luo, L. (2017, December). Application of Data Mining in Library-Based Personalized Learning. International Journal of Emerging Technologies in Learning (iJET), 12(12), 127–133. https://doi.org/10.3991/ijet.v12i12.7967.
- L. Zheng& J. Guo& W. Wang & R. Wang. (2019). Research on Bibliographic Push Service of University Library Based on Data Mining. In 3rd International Conference on Data Science and Business Analytics (ICDSBA), Turkey. 288-290. https://081040nvo-1106-y-https-ieeexplore-ieee-org.mplbci.ekb.eg/document/9270175.
- Martínez-Martínez, J.M., Escandell-Montero, P., Soria-Olivas, E. (2016, January). A new visualization tool for data mining techniques. Prog Arif Intel 5, 137–154. https://doi.org/10.1007/s13748-015-0079-4.
- Martínez Abad, F., & Chaparro Caso López, A. A. (2016, September). Data-mining techniques in detecting factors linked to academic achievement. School Effectiveness and School Improvement, 28(1), 39–55. https://doi.org/10.1080/09243453.2016.1235591.

- Nandha Kumar, K.G., Christopher, T. (2015). Applications of data mining techniques in academic libraries. International Journal of Applied Engineering Research, 10(55), 1500–1502.https://0810511ws-1106-y-https-www-scopus-com.mplbci.ekb.eg/record/display.uri?eid=2-s2.0-84942435226&origin=resultslist&sort=plf-f&src=s&sid=51420b12dba668a04f89f020c913141a&sot=b&sdt=b&s=TITLE%28data+AND+mining+AND+libraries%29&sl=40&sessionSearchId=51420b12dba668a04f89f020c913141a&relpos=90.
- Nivedhitha G& Rupavathy N. (2018, February). Data mining in personalized service of digital libraries. International Journal of Engineering and Technology (UAE),7(17), 51 – 53. https://doi.org/10.14419/ijet.v7i1.7.9570
- Pang, Lu. (2022, September). Library Management System Based on Data Mining and Clustering Algorithm. Wireless Communications and Mobile Computing. https://doi.org/10.1155/2022/1398681
- Qiao, L. (2022, July). Application of Big Data Mining Technology in the Digital Construction of Vocal Music Teaching Resource Library. Web Intelligence, 20(2), 123 - 131. https://doi.org/10.1155/2022/3197118.
- Qin Jianning. (2023). The Use of Data Mining in Long Tail Service of Library. In 9th Annual International Conference on Network and Information Systems for Computers. https://doi.org/10.1109/ICNISC60562.2023.00055
- R. Ochilbek. (2019). Using data mining techniques to predict and detect important features for book borrowing rate in academic libraries. In the 15th International Conference on Electronics, Computer and Computation (ICECCO), 1-5. https://doi.org/10.1109/ICNISC60562.2023.00055
- Reitz, J. M. (2010). ODLIS: Online Dictionary for Library and Information Science. Available at: http://www.abc-clio.com/ODLIS/odlis_c.aspx.
- Semaan, Paul. (2012, June). Natural Language Generation: An Overview. Journal of Computer Science & Research (JCSCR).1(3), 50-57. http://www.lacsc.org/papers/PaperA6.pdf.
- Shahnaz Khademizadeh& Zahra Nematollahi& Farshid Danesh. (2022, October). Analysis of book circulation data and a book recommendation system in academic libraries using data mining techniques. Library & Information Science Research. 44(4). https://081010n7x-1106-y-https-doi-org.mplbci.ekb.eg/10.1016/j.lisr.2022.101191
- Siguenza-Guzman, L.& Saquicela, V.& Avila-Ordóñez, E.& Vandewalle, J.& Cattrysse, D. (2015, July). Literature Review of Data Mining Applications in Academic Libraries. Journal of Academic Librarianship, 41(4), 499–510. https://doi.org/10.1016/j.acalib.2015.06.007
- Simões., Marques, M., Figueira, J. (2018). How Can AI Help Reduce the Burden of Disaster Management Decision- Making?. In the International Conference on Applied Human Factors and Ergonomics. Springer. https://doi.org/10.1007/978-3-319-94334-3
- Surya Rebbapragada, Amit Basu, John Semple. (2010, April). Data mining and revenue management methodologies in college admissions. Commun,53(4),128–133. https://doi.org/10.1145/1721654.1721690.
- T. Zhu, L. Zhang. (2011). Application of data mining in the analysis of needs of university library users. In the 6th International Conference on Computer Science & Education (ICCSE), 391-394.https://0810412n0-1106-y-https-doi-org.mplbci.ekb.eg/10.1109/ICCSE.2011.6028662.
- Wei, Yanga& Machica, Ivy Kim D& Arroyo, Jan Carlo T& Sabayle, Ma Luche P& Delima, Allemar Jhone P. (2022, May). Analysis of Students' Borrowing Behavior in Local University Library Using Data Mining Technique. International Journal of Emerging Technology and Advanced Engineering, 12(7), 68– 77. https:// DOI: 10.46338/ijetae0722_07.

- Weifeng, Wang& Luyang, Meng& Lei, Wu& Jing, Zhang. (2020, May). Research and Application of Data Mining Technology in Library Office Information Construction. Journal of Physics: Conference Series. https:// doi:10.1088/1742-6596/1550/3/032001.
- Xiao Lin& Wenjuan Guan& Ying, Zhang. (2023, Novembre). Application of Data Mining Technology with Improved Clustering Algorithm in Library Personalized Book Recommendation System. International Journal of Advanced Computer Science and Applications (IJACSA), 14(11). http://dx.doi.org/10.14569/IJACSA.2023.0141151
- Xiao, Q. (2020, September). Resource classification and knowledge aggregation of library and information based on data mining. Ingénierie des Systèmes d'Information, 25(5), 645-653. https://doi.org/10.18280/isi.250512.
- Xu, Chunyan& Bai, Jing. (2022, June). Massive-Scale Data Mining to Enhance Digital Library with Applications in College Education. Applied Bionics and Biomechanics, 7. https://doi.org/10.1155/2022/9698477
- Yanping Cheng& Qingyu Liu.(2019). Application of data mining in the evaluation system of University library service. In IEEE 4th International Conference on Big Data Analytics (ICBDA), 99-10. 10.1109/ICBDA.2019.8713202
- Y. Cheng, Q. Liu. (2019). Process and Application of Data Mining in the University Library. In 4th International Conference on Big Data Analytics (ICBDA), 123-127. Available at: https:// doi: 10.1109/ICBDA.2019.8713202.
- Y. M. Tashtoush& M. Al-Soud& M. Fraihat& W. Al-Sarayrah & M. A. Alsmirat. (2017). Adaptive e-learning web-based English tutor using data mining techniques and Jackson's learning styles. In 8th International Conference on Information and Communication Systems (ICICS). Jordan, 86-91. https:// doi: 10.1109/IACS.2017.7921951.
- You, S., Joo, S. and Katsurai, M. (2024, January). Data mining topics in the discipline of library and information science: analysis of influential terms and Dirichlet multinomial regression topic model. Aslib Journal of Information Management, 76(1), 65-85. https://0810bxenf-1103-y-https-doi-org.mplbci.ekb.eg/10.1108/AJIM-05-2022-0260
- Yu.(2011).Data Mining in Library Reader Management. in International Conference on Network Computing and Information Security, 54-57. https:// doi: 10.1109/NCIS.2011.109.
- Zhang Jing. (2021). Application of Data Mining Based on Computer Algorithm in Personalized Recommendation Service of University Smart Library. Journal of Physics: Conference Series. Available at: https:// doi:10.1088/1742-6596/1955/1/012008.
- Zong, Y. (2023). Personalized Recommendation Service of University Library Based on Data Mining Technology. In the 2nd International Conference on Cognitive Based Information Processing and Applications (CIPA 2022). CIPA 2022, (55). https://doi.org/10.1007/978-981-19-9373-2_84

Applying of Data Mining Mechanisms on Improving the Level of Circulation Service: An Applied Study on Adult Circulation Data at the Public Library of Egypt

Dr. Kareman Baknam Sedki

Lecturer at Library, Archive & Information Technology, Dep.
Faculty of Arts, Cairo University
Kareman14@cu.edu.eg

Abstract:

In light of the technological developments witnessed in the era of information technology, the inflation and diversity of information sources, and the massive data they provide that is difficult to handle, there has emerged a need to develop more capable methods for conducting research and data mining as a technique aimed at deriving knowledge from vast amounts of data, transforming it from accumulated and incomprehensible data into valuable information that can be utilized to build decision making strategies.

From this perspective, this study aims to invest in data mining methods to analyze adult circulation files at the Public Library of Egypt. The study relies on a descriptive approach using analytical methods to present the main concepts related to data mining, its applications, fundamental issues, and tasks, as well as the methods and techniques used in data mining, clarifying the dynamics of each. Additionally, it employs open-source technologies, including the Orange3 program, to extract knowledge from the adult circulation files at the Public Library of Egypt, and how to benefit from that knowledge in revealing their research needs and reading preferences, and providing suggestions for supplying them with information sources that meet those needs, in addition to deriving new insights for the development of the library's collections.

One of the most important conclusions of this study is that libraries and information centers should be directed towards utilizing data mining techniques to provide their services with high efficiency and quality. This is due to the information it provides about the users and the identification of their reading needs, which contributes to offering a new perspective on the information sources they possess and the services they provide to meet the needs of its users. Additionally, data mining tools can be utilized to build a database that assists library staff and decision-makers in making informed decisions regarding the development of library collections and the future services that should be offered.

Keywords: ***Data Mining; Artificial Intelligence; Egypt Public Library; Reading and Research Needs; Orange3 Program; Natural Language Processing.***