أهمية تنقيب البيانات في كشف الاحتيال في القوام المالية

الأستاذ المساعد الدكتور سيد مهدي معلمي (الكاتب المسؤول)
أستاذ في العلوم الاقتصادية في جامعة المصطفى ﷺ في قم وطهران ومشهد
Mahdimoalemi@gmail.com
الباحث جمال نعمان محمد الحلو
مسؤول مفقودات مطار النجف الاشرف الدولي
jamalnajafairport@yahoo.com

The importance of data exploration in detecting fraud in financial statements

Associate Professor Dr. Sayed Mahdi Moalemi (Responsible author)
Professor of Economics at Mustafa University in Qom, Tehran and Mashhad
Researcher Jamal Noaman Mohammed Al-Hilo
Missing Persons Officer of Najaf Ashraf International Airport

الملخص:

تمر البشرية في مرحلة تفوق العلوم والتقنيات، ولا يخفى على أحد الأثر الذي تركه هذا التفوق في حياة الإنسان، ومما لا شك فيه أن المجتمعات التي استملكت المعلومة والمعرفة هي الآن التي تحكم العالم وتتصدر مشهد الدول المتقدمة والعصرية.

أحد أشكال التقدم والتقنية وأحد أبرز آثارها هو تضخم المعلومات الهائل، والذي يحدث باضطراد مستمر، حيث أصبحت مخازن البيانات وقواعد البيانات التي تعتمد عليها أغلب النظم والتطبيقات، أصبحت ممتلئة بكم كبير وضخم من البيانات العشوائية والتاريخية، ونستطيع القول فعلاً أن البيانات لا تنام ولا تتوقف عن النمو.

مع هذا النمو في البيانات، أصبحنا أمام تحدي كبير يتمثل في كيفية الاستفادة من هذه البيانات المُخزنة في مستودعاتها، وكيفية القيام بعملية استخراج المعلومة المفيدة منها، وهذا التحدي أصبح لا يمكن حله بواسطة الطرق التقليدية لتحليل البيانات، لذلك فإننا بحاجة إلى تقنيات غير تقليدية تستوعب التعامل مع الكميات الضخمة للبيانات بالإضافة للقدرة على التعامل مع الأنواع والاشكال المختلفة للبيانات، وهذا يعني تقنيات على مستوى عالي من الكفاءة والذكاء.

الكلمات المفتاحية: تنقيب البيانات، الاحتيال، القوائم المالية.

Abstract:-

Humanity is at the stage of superiority of science and technology, and no one is aware of the impact of this superiority in human life, and it is no doubt that the societies that have acquired information and knowledge are now that govern the world and top the scene of developed and modern countries.

One of the most prominent forms of progress and technology, one of the most prominent effects is the huge and constantly increasing information overload. Data stores and databases on which most systems and applications rely are now full of large random and historical data, and we can say that data never sleeps. the growth.

With this growth in data, we are faced with a big challenge of how to take advantage of the data stored in its repositories and how to extract useful information from it. This challenge can not be solved by traditional methods of data analysis. With huge amounts of data in addition to the ability to deal with different types and types of data, which means technologies at a high level of efficiency and intelligence.

Keywords: Data exploration, Fraud, Financial statement.

المقدمة:ـ

أدى الانتشار الواسع لتقنية المعلومات وسهولة إتاحتها إلى تضخم حجم المعلومات بصورة استباقية لم يشهدها التاريخ من قبل، مما جعل من قضية البيانات الضخمة على الإنترنت مثاراً للجدل، من حيث جدوى وجودها بهذه الصورة العشوائية.

وعندما نتحدث عن البيانات الضخمة، فإننا نتحدث عن كميات لا يمكن تخيلها من البيانات متعددة الأنواع والمصادر بحجم يصل إلى المئات من التيرابايت أو حتى البيتابايت (البيتابايت هو الرقم واحد متبوعا بـ ١٥ صفر).

من هنا ظهر ما يسمى باستخراج البيانات Data Mining كتقنية تهدف إلى استنتاج المعرفة من كميات هائلة من البيانات، تعتمد على الخوارزميات الرياضية والتي تعتبر أساس التنقيب عن البيانات وهي مستمدة من العديد من العلوم مثل علم الإحصاء والرياضيات والمنطق وعلم التعلم، والذكاء الاصطناعي والنظم الخبيرة، وعلم التعرف على الأنماط،وعلم الآلة. وغيرها من العلوم والتي تعتبر من العلوم الذكية وغير التقليدية.

ظهر التنقيب في البيانات (Data mining) في أواخر الثمانيات وأثبت وجوده كأحد الحلول الناجحة لتحليل كميات ضخمة من البيانات، وذلك بتحويلها من مجرد معلومات متراكمة وغير مفهومة (بيانات) إلى معلومات قيِّمة يمكن استغلالها والاستفادة منها بعد ذلك.

وقد اجتذبت مرحلة التنقيب في البيانات الكثير من الاهتمام في الأوساط البحثية على مدي العقد الماضي، في محاولة لتطوير خوارزميات قابلة للتوسع والتكيف مع كميات متزايدة من البيانات في البحث عن أنماط معرفية ذات معنى.

وقد نمت حزم من الخوارزميات والبرمجيات و بشكل كبير خلال العقد الماضي، إلى حد أن التوسع قد جعل من الصعب على العاملين في هذا الحقل تتبع التقنيات المتاحة لحل مهمة معينة.

الفصل الأول

منهجية البحث

١-١-١ أهمية الدراسة The importance of the study

أهمية دراسة تقنيات التنقيب عن البيانات في إدارة العمليات المصرفية والمحاسبية في البنوك، اعتقادا من الباحثين في أن أهمية التنقيب عن البيانات بمثل أهمية البيانات والمعلومات اللازمة لتنفيذ العمليات المصرفية والمحاسبية في البنوك و ايجاد الطية بين عناصر العمليات المصرفية.

١-١-٢ مشكلة الدراسة The problem of the study

ويمكن تلخيص مشكلة الدراسة من خلال طرح التساؤلات التالية:

١- هل أن وجود كم هائل من البيانات والمعلومات المحاسبية و غير المحاسبية المخزنة أو المسترجعة أو المعاد استخدامها فبي البنوك، يتطلب استخدام تقنيات التنقيب عن البيانات لتحقيق إدارة العمليات المحاسبية والمصرفية في البنوك؟

٢- هل أن استخدام تقنيات التنقيب عبن البيانات في البنوك الاردنية تؤدي إلى كفاءة إدارة العمليات المحاسبة والمصرفية في البنوك.

3-١-١ أهداف الدراسة The objectives of the study

هدفت الدراسة الى بيان مستوى الاهتمام بتطبيق وبالمجالات المكونة لبيئة مفاهيم وادوات التنقيب عن البيانات لإدارة العمليات المصرفية والمحاسبية في البنوك التجارية الاردنية. ولتحقيق تلك الاهداف استخدام الباحثان المنهج الوصفي التحليلي القائم على الاستبيان الموزع على افراد مجتمع الدراسة. وقد توصل الباحثان الى أن نسبة الاهتمام لدى أفراد مجتمع الدراسة عن تطبيق مفاهيم التنقيب عن البيانات لإدارة العمليات المحاسبية والمصرفية، كانت مرتفعة بشكل عام و أن ترتيب المجالات التي قد ينظر إليها عند البحث في تطبيق مفاهيم التنقيب عن البيانات والتي تناولتها هذه الدراسة كانت بحسب اهميتها ومستوى الاهتمام بها لدى أفراد مجتمع.

٤-١-١ فرضية الدراسة The hypotheses of the study

إن أهمية تنقيب البيانات Data mining في كشف الاحتيال في القوام المالية هي للتحقق من كفاية وفاعلية الاستخدام وحمايته الأموال من سوء الاستخدام.

الفصل الأول

تقنية تنقيب البيانات: Data Mining Technology

هي تكنولوجيا جديدة تهدف الى الحصول على معلومات غير معروفة من قواعد البيانات لغرض استخدامها في عملية اتخاذ القرار،كما انها تحسن من فهم البيانات المتراكمة واستخلاص ثروة المعلومات الموجودة فيها من اجل استثمارها بالشكل الذي يؤدي الى اتخاذ قرار سليم يجنب الوحدة الاقتصادية اي خسائر محتملة ويرفع من مستوى ادائها.

مراحل اكتشاف المعرفة Stages of knowledge discovery **:**

اكتشاف المعرفة في قواعد البيانات Knowledge Discovery in Database (KDD) ليس بالعملية السهلة والتي قد يعتقد البعض أنها تتوقف عند تجميع البيانات وإدارتها، بل نراها تمتد إلى التحليل والتوقع والتنبؤ بما سيحدث مستقبلاً.

التنقيب في البيانات يشكل جزءا من اكتشاف المعرفة knowledge discovery، وهذه العملية هي الأكثر شمولا. تتضمن عملية اكتشاف المعرفة الخطوات التالية:

١- اكتشاف البيانات Data discovery: وهي مرحلة جمع البيانات وتشمل كشف وتحديد وتوصيف البيانات المتاحة.

٢- تصفية البيانات وتنقيتها Data cleaning: ويتم في هذه المرحلة إزالة البيانات المزعجة Noise التي لا أهمية لها، كما يتم حذف البيانات المتضاربة والبيانات الغير متناسقة.

٣- تكامل البيانات Data integration: يتم في هذه المرحلة تجميع البيانات المتشابهة وذات الصلة من مصادر البيانات المتعددة ودمجها معا.

٤- اختيار البيانات Data selection: في هذه المرحلة، يتم تحديد واسترجاع البيانات الملائمة من مجموعة البيانات.

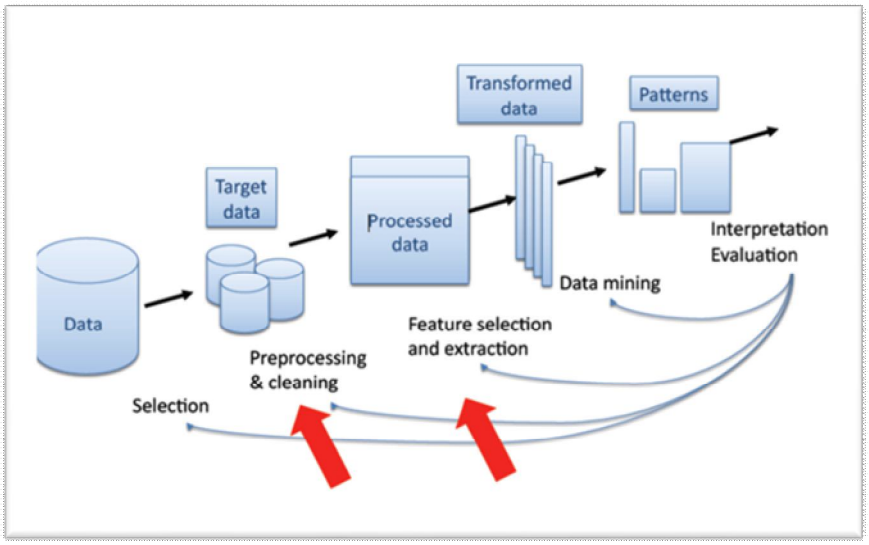
٥- تحويل البيانات Data transformation: في هذه المرحلة يتم تحويل البيانات إلى نماذج مخصصة ملائمة لإجراءات البحث والاسترجاع بواسطة خلاصة الإنجاز أو عمليات التجميع.

٦- التنقيب عن البيانات Data mining: أي استخدام طرق ذكية تطبق لاستخلاص أنماط البيانات استخراج نماذج مفيدة قدر الإمكان.

٧- تقييم النمط Pattern evaluation: يتم في هذه المرحلة تحديد الأنماط المهمة حقا والتي تمثل قاعدة المعرفة لاستخدام بعض المقاييس المهمة.

٨- تمثيل المعرفة وتقديمها Knowledge presentation: وهي المرحلة الأخيرة من مراحل اكتشاف المعرفة في قواعد البيانات وهي المرحلة التي يراها المستفيد، هذه المرحلة الأساسية تستخدم الأسلوب المرئي لمساعدة المستفيد في فهم و وتفسير نتائج استخراج البيانات.

Data Mining Model

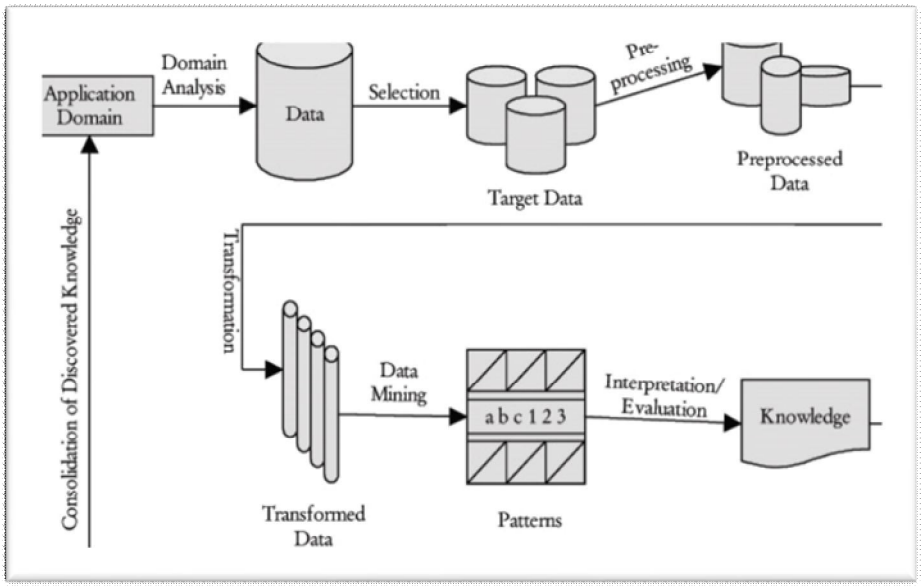


وبالتالي التنقيب في البيانات هو خطوة أساسية لتطبيق أساليب ذكية بهدف الكشف عن أنماط البيانات المثيرة للاهتمام والمخبأة في مجموعات البيانات الكبيرة. ومع ذلك، في بعض المنظمات نجد أن مصطلح التنقيب عن البيانات "data mining" أصبح أكثر شعبية للإشارة

إلى العملية التي يتم فيها اكتشاف المعرفة knowledge discovery برمتها.

وهناك جانب مهم جدا، يجب النظر إليه بعين الاعتبار و هو أن هناك أنماط جديدة قد تبرز،عادة ما تكون غير معروفة من قبل. وبالتالي يجب أن تكون أدوات التنقيب عن البيانات قادرة على البحث عن أنواع مختلفة من الأنماط، بأشكال متوازية لزيادة كفاءة التنقيب عن البيانات. كما يجب أيضا أن يتم الكشف عن الأنماط في الأجزاء الصغيرة والفرعية مما يعني البحث في مستويات مختلفة من التجريد أو التفصيل. حلول التنقيب عن البيانات الجيدة هي التي تشير أيضا إلى قدر من الثقة أو اليقين المرتبطة مع نمط اكتشافها، لأن بعض الأنماط قد لا تصلح لكافة البيانات في مجموعة البيانات التي تم تحليلها.

Knowledge discovery by exploration data model



اصبح دور نظم المعلومات في هذه البيئة الجديدة ممارسة الدور المعرفي كنظم متكاملة تضيف إلى أدوارها التقليدية أدواراً جديدة تعمل على توفير معلومات ومعارف واسعة، تُساعد في تحليل البيانات المالية بشكل أفضل، ويتناول هذا الدور ما يأتي.

١- إيجاد المعرفة وتأمينها (Creating Knowledge)

تسعى نظم البحث للمعلومات المحاسبية في تجهيز العاملين في الحقل المعرفي بالرسومات، والتحليلات، والاتصالات، ووسائل إدارة الوثائق، فضلاً عن نظم البحث والاسترجاع للوصول إلى مصادر المعلومات والمعرفة الداخلية والخارجية.

١- اكتشاف وتصنيف المعرفة (Discovering and Codifying Knowledge)

وفرت تقنيات التنقيب عن البيانات إمكانية استنباط الخبرات ودمجها، لغرض إيجاد نماذج وعلاقات، في كميات كبيرة من البيانات، وتقوم نظم البحث والاسترجاع بتحليل قواعد بيانات واسعة، وتستطيع أيضاً اكتشاف معارف جديدة منها.

٢- المشاركة بالمعرفة (Sharing Knowledge)

٤- فنظم البحث للمعلومات المحاسبية التي توفرها آليات تنقيب التعاون الجماعية تستطيع أن تُساعد العاملين على الوصول والعمل في آن واحد على الوثيقة نفسها، ومن مواقع مختلفة، ومن ثم التنسيق بين نشاطاتهم المختلفة.

٣- توزيع المعرفة (Distributing Knowledge)

فنظم البحث لتقنيات التنقيب عن البيانات وأدوات الاتصال الخاصة بها تستطيع تأمين الوثائق والأشكال الأخرى من المعلومات، وتوزيعها على العاملين في مجال المعلومات والمعرفة، بغرض ربط المكاتب بوحدات الأعمال الأخرى داخل المنظمة وخارجها. ومن ثم فإن المميزات التي توافرها تقنيات التنقيب في البيانات في توفير معلومات دقيقة، وصحيحة، وبشكل سريع، تبين أهمية هذه التقنية أهم هذه المميزات:

- تُسهل التعامل مع تقنيات المعلومات المتطورة، وتُساعد على قياس فاعلية نُظم.
- تُساعد على الاستخدام الفعال لمصادر البيانات والموارد المتاحة، وفي التدقيق لتحسين نُظم المعلومات المحاسبية المستخدمة.
- تساعد على الكشف عن قدرة المنشأة على النمو ومواكبة التطور، وحاجتها إلى تطوير المصادر الفنية والبشرية لنُظم المعلومات المستخدمة، ومن ثم فإن المعلومات تُساعد على زيادة المعرفة، والحد من البدائل، وهذا الأمر يؤدي إلى التخلص من حالة عدم التأكد التي تتمثل في هذه البدائل.

- تُمكن المدقق في أن يخطط لاستمرارية إجراء الفحص وتحديد المشكلة وعناصرها وإتمام الرقابة على ملفات المعلومات والبيانات اللازمة لقراراته، ويسعى إلى تطوير نظم المعلومات.

تصنيف منهج تنقيب البيانات Classification of data mining approach

١. التصنيف Classification:يتم تصنيف البيانات بناء على الخصائص المشتركة لها، والهدف من التنصيف بناء نموذج لتصنيف كل مجموعة صفات الى صف واحد في مجموعة الصفوف وتستخدم شجرة القرار (Nearest Neighbor Algorithm) الأقرب الجار خوارزمية التي تهدف الى التنبؤ عن طريق مقارنة السجلات الشبيهة بالسجل المراد التنبؤ له وتقدير القيمة المجهولة لهذا السجل بناء على معلومات لتلك السجلات والشبكات العصبية.

٢. الانحدار Regression: هو تقنية تسمح بتحليل البيانات لوصف العلاقة بين متغيرين أو أكثر، إن الانحدار يفترض أن توضع البيانات بنوع معروف من الدوال ومن ثم يتم تحديد أفضل دالة للبيانات المعطاة.

٣. تحليل السلاسل الزمنية Time Series Analysis: إن مشاهدة البيانات عبر الزمن تنتج تحليلاً مفيداً لأنه تتم مشاهدة سلوك البيانات عبر الزمن بشكل أساسي. وهذا يعني أن قيم الصفة المميزة للبيانات التي يتم فحصها تكون متغيرة عبر الزمن

٤. التنبؤ Prediction: يعد التنبؤ من الأدوات التي تجذب الانتباه لأنها تتمكن من إعطاء مغزى التوقع الناجح في سياق العمل لذا فإنه يمكن النظر إلى العديد من تطبيقات تنقيب بيانات العالم

الحقيقي كأنها تنبؤ بحالة بيانات مستقبلية معتمدة على بيانات سابقة وحالية.

٥-التجميع Clustering: ويمكن تعريفها على أنها وصف الخصائص العامة للنماذج وتنظيم البيانات في اصناف.

٦- المرئية Visualization: إن مرئية النتائج يكون مساعداً في تسهيل ملاحظات مخرجات خوارزميات تنقيب البيانات وفهمها.

تطبيقات تنقيب البيانات Data Mining Applications

امتدت تطبيقات تنقيب البيانات الى مجالات عديدة وواسعة ولا زالت مستمرة في الدخول إلى مجالات جديدة تضاف الى ما سبق وهي كالآتي:

(Avison & Shah , 1997:328) (Ramachandran , 2001: 3) (Wu , 2002:2)

(Tow Crows , 2002:1)

١- الأعمال المصرفية Banking كتحليل مخاطر القروض

٢- المالية Financial كتحديد الغش في تبادل الاسهم

٣- الاتصالات Telecommunicationsكتحديد الاستخدام المزيف لخدماتها

٤- التسويق Marketing كايجاد العلاقة بين الخصائص الديموغرافية للزبون

٥- التامين ورعاية الصحة Insurance and Health Care كتحليل الدعاوى

٦- الطب Medicine كتحديد العلاجات الطبية الناجحة للأمراض المختلفة

٧-النقل Transportation كتحديد جدول التوزيع بين المنافذ

٨-البيع بالتجزئة Retailing كتقدير فاعلية الترويج

٩- إدارة علاقات الزبون للاحتفاظ بالزبون لانه اقل كلفة من كسب زبون جديد Customer Relationship Management كاتخاذ الفعل اللازم

١٠-رقابة الجودة او تحليل الخطأ الخصائص المحيطة بالمنتجات المتضررة ومن ثم تحسين الجودة من خلال اجراء التغيراتQuality Control or Error Analysis كتحديد كتحديد خصائص أفراد الأداء العالي

١١-التجارة الإلكترونية Hiring

١٢-تحليل الضمان Warranty Analysis

تصنيف تقنيات استخراج البيانات لكشف الغش المحاسبة المالية:

Classification of data mining techniques to detect fraud and financial accounting

لتحديد خوارزميات الرئيسية المستخدمة للكشف عن الاحتيال في المحاسبة المالية، يقدم استعراض تقنيات تنقيب البيانات تطبيقها على الكشف عن الاحتيال المالي. أكثر التقنيات استخداما هي نماذج اللوجستية (الانحدار)، والشبكات العصبية، وشبكة بايس العقائدية (Bayesian Belief Networks) تطبيقات الذكاء الاصطناعي وأشجار القرار، وكلها تندرج في فئة التصنيف. وتناقش هذه التقنيات أربعة بمزيد من التفصيل في الفقرات التالية.

١- **نموذج الانحدار:** regression model

استمرت عملية التدقيق في تطور متأثرة بمتغيرات عديدة منها طبيعة النظام الاقتصادي وأنواع الملكية وأحجام الشركات وتعدد عملياتها ومدى تعقيدها، فأصبح الأمر يتطلب عملية التحقيق بالإضافة إلى الفحص. مما يتطلب من المدقق الخارجي استخدام تقنيات حديثة تساعده في عمليات اتخاذ القرار المناسب ومن هذا التقنيات استخدام تحليل لانحدار، وقد اقترحت النموذج للتحقيق في العلاقة بين التداول من الداخل وإمكانية الاحتيال. الدراسة في وجدت أن، يتم تنفيذ عملية الاحتيال من قبل كبار المسؤولين التنفيذيين والمديرين، من خلال الحد من حيازات الأسهم من خلال ارتفاع النشاط بيع الأسهم. أساليب أخرى مثل تحليل الانحدار الإحصائي مفيدة أيضا لاختبار إذا كان وجود لجنة مستقلة لمراجعة الحسابات يخفف أو يقلل من احتمالات التزوير. كما أن المنظمات التي لديها لجان المراجعة، التي شكلتها من مدراء مستقلين، والقيام بالتدقيق لا يزيد عن مرتين في السنة، هي أقل عرضة للاحتيال في اعداد التقارير المالية. تحليل الانحدار باستخدام نموذج اللوغارتميين يمكن استخدامها لتحليل تجريبي للمؤشرات المالية التي يمكن التنبؤ الاحتيال المالي بشكل كبير. تحليل وتجميع اللوجستي تحليل مشترك يمكن استخدامها لوضع نموذج كشف الاحتيال من أربعة جوانب من المؤشرات المالية والحوكمة شركات، المخاطر المالية والضغوط والتجارية ذات الصلة.

٢- **الشبكات العصبية** Neural networks:

وهي قريبة من اشجار القرار لكنها اصعب فهما" منها ولا توضح نتائجها وتستغرق وقتا" أطول وتقدم نماذج ذات قوة تنبؤية افضل. حيث متغير تتكون الشبكة العصبية من طبقات تبدأ بطبقة المدخلات Input Layer التي ترتبط بدورها التنبؤ وترتبط عقد المدخلات بعدد من عقد الطبقة الخفية Hidden Layer حيث تضم واحد او اكثر من بعقد طبقة خفية اخرى او بطبقة المخرجات Output Layer

٣- **شجرة القرار** (Decision Trees(DT:

وهي هيكل شجرة يتكون من عقد تمثل العقدة اختيار على بديل (ميزة) وكل فرع يمثل ناتج اختيار البديل وبهذه الحالة تمثل الشجرة محاولات لتقسيم المشاهدات الى مجموعات فرعية يستبعد بعضها بعضا،ويتم اختيار البديل الافضل، وتقسم الفروع الى فروع اصغر الى حدود عدم الجدوى من التقسيم. وهناك العديد من خوارزميات التقسيم المقترحة مثل AID او استخدام اختبار t كما في تحليل التباين. ويستخدم AID مربع كاي – chi square، كلها يمكن استخدامها في تصنيف الاشجار، ويستخدم تحليل الانحدار كمؤشر للتباين.

٤- **المنطق الضبابي والخوارزميات الجينية**Fuzzy logic and Genetic Algorithm

تستخدم الخوارزميات الجينية في أنظمة تصنيف لتمثيل ونمذجة سلوك قرار المدقق في بيئة الاحتيال تعتبر الخوارزميات الجينية من التقنيات الهامة في البحث عن الخيار الأمثل من مجموعة حلول متوفرة لتصميم معين، وتعتمد مبدأ داروين في الاصطفاء حيث تقوم هذه المعالجة الوراثية بتمرير المزايا المثلى من خلال عمليات التوالد المتعاقبة.

٥- **الأنظمة الخبيرة:** Expert Systems

الباحثون قاموا بفحص دور النظم الخبيرة في زيادة قدرة مدققي الحسابات والمستخدمين في الكشف عن الاحتيال. باستخدام نظام خبير، فإنها يمكن أن يكون الافضل للكشف عن مخاطر الاحتيال المحاسبي وبالتالي يساعد المدققين في إعطاء اقتراحات التدقيق موثوقة بكثير من خلال إجراء التدقيق العقلاني. وأكدت الأبحاث أن استخدام نظام خبير تعزيز اداء المدققين مع مساعدة من نظام خبير، نظام خبير ساعد في صنع القرار بشأن الإجراءات التدقيق المناسبة.

ووفقا للتطبيق استخراج البيانات وتقنية تنقيب البيانات. وقد حاول بعض الباحثين لتطبيق العديد من مجموعات تقنيات التنقيب البيانات مثل تحليل النسب ، اشجار القرار، الشبكات العصبية، والهدف الرئيسي هو كشف عن البيانات المالية الاحتيالية.

نموذج عملي لتقنية تنقيب البيانات في الفحص الاحتيالي

Practical model for data mining technology in the judicial examination

تتمزٌ تقنيةٌ التنقبٌ في البيٌنات عن التقنيات التنقيب الأخرى في أن التقنيٌات الأخرى تحاول تحدد الشواذ من خلال استعمال مجموعة من قواعد الترشيحٌ على متغيرٌ معينٌ، في حينٌ إن تقنيةٌ التنقيبٌ عن البيٌانات تعتمد على محاولة اكتشاف العلاقة الموجودة بينٌ عدة متغيرٌات. ومن الامور الأخرى التي يجب أن تؤخذ بنظر الاعتبار هيً إن تنقيبٌ البيٌانات يعٌد مجرد احد عناصر عملية التحري للمحاسبة القضائيةٌ، وان عمليةٌ التحري لا يٌمكن أن تكتمل من خلال شاشة الكومبيوٌتر فقط وإنما تتطلب عمليةٌ التحري مراجعة المستندات وإجراء المقابلات وغيرها من أعمال التحري. كما تٌطلب الأمر الأخذ بنظر الاعتبار ضرورة التحقق من دقة البيٌانات التي تٌم الحصول عليها وكذلك التحقق من اكتمالها، فضلا عن ضرورة الأخذ بنظر الاعتبار المسائل القانونيةٌ وبعناية كبيرٌة إذ إن البيٌئة القانونيةٌ ليسٌت متشابهة في جميعٌ البلدان لذلك يجب تؤخي الحذر قبل البدء بعمليةٌ جمع البيٌنات وتحليلٌها والتأكد من إن الإجراءات المخططة مسموح بها من وجهة النظر القانونيةٌ

(Golden and others 2006, 388)

تعد هذه التقنيةٌ من التقنيٌات الأخرى المفيدٌة في اكتشاف الاحتيٌال من خلال احتساب نسب تحليلٌ البيانات للحقول الرقميةٌ الرئيسيٌةٌ. فكما هو الحال بالنسبة للنسب الماليةٌ التي تعطي مؤشرات عن الوضع المالي للشركة، فإن نسب تحليلٌ البيٌنات تبلغ عن وضع الاحتيٌال من خلال تحديدٌ الأعراض الممكنة للاحتيٌال. وهنالك ثلاثة نسب تستعمل في هذا المجال هي

١- نسبة أعلى قيمٌة إلى أدنى قيمٌة (max/mine)

٢- نسبة أعلى قيمٌة إلى ثان أعلى قيمٌة (max/max2)

٣- نسبة العام الحال إلى العام السابق -

فمثلا من خلال استعمال تحللٌ النسب يقوم المحاسب القضائي بدراسة العلاقات بينٌ تكاليف محددة وبعض مقاييسٌ الإنتاج، مثل الوحدات المباعة، إيرٌاد المبيعٌات أو ساعات العمل المباشر. على سبلٌ المثال، للوصول إلى التكاليف الغير المباشرة لكل ساعة عمل مباشر مٌكن تقسيمٌ التكاليف الغير المباشرة على إجمالي -ساعات العمل المباشر. كما إن تحليلٌ النسب قد يسٌاعد المحاسب القضائي في تقديرٌ بعض المصاريف على سبيلٌ المثال، إذا كانت كلفة الوحدة المباعة من احدى سلع الشركة مٌكن تقد رٌها بنسبة ٣٥٪ من ايرٌاد المبيعٌات، عندها مٌكننا استعمال هذه النسبة لتحديدٌ مبلغ معقول لكلفة البضاعة المباعة عند فقدان السجلات المالية للشركة قيدٌ التدقيق (Metha and Mathur, 2007, 1577)

إن استعمال النسب السابقة قد يودي إلى اكتشاف الأخطاء أو عملياٌت الاحتياٌل في السجلات، إذ أوضحت عمليةٌ التدقيقٌ وباستعمال النسب إلى تحديدٌ الأخطاء التي حدثت في المدفوعات إلى المجهزون وكما يلي:

المجهزين	اكبر دفعة	ثاني اكبر دفعة	نسبة اكبر الى ثاني اكبر دفعة	اصغر دفعة	انسبة اكبر الى اصغر دفعة	عدد الدفعات
ABC Grop	310,345,21	5,441,62	20.28	123,93	890,38	314
XYZ LTd	23,489,92	3,641,97	6,79	749,37	31,35	763
ZZZ CO	299,345,00	4,281,03	69,72	442,28	676,38	140

Source: (Coderre, 2002, 16)

تٌبينٌ من البيٌانات أعلاه بأن هنالك تباينٌ كبيرٌ بينٌ اكبر دفعة وثان اكبر دفعة لكل من الموردينٌ الثلاثة، وان ذلك جٌذب انتباه أو اهتمام المحاسب القضائي لكون اكبر دفعة كانت بعيدٌة جدا عن الحدود المعتادة لها وان عدم وجود تفسيرٌ لذلك سيؤٌدي إلى ارتفاع نسبة الشك بوجود الاحتياٌل، وعند التحري عن الدفعات المسددة إلى المورد ABC Corp وجد بأن هنالك ٣١٣ دفعة كانت تقع بينٌ ١٢٣.٩٣ دولار و ٥٤٤١.٦٢ دولار وان هنالك دفعة واحدة فقط بمبلغ ١١٠٣٤٥.٢١ دولار والتي تزيدٌ بأكثر من ٢٠ ضعف مقارنة بثان اكبر دفعة، وعند التحقيقٌ تبينٌ بأن مديرٌ المشتريٌات)الذي قدم استقالته بعد أسبوعينٌ من تاريخٌ الدفعة (قد عقد صفقة مشتريات بذلك المبلغ مع الشركة التي عرضت علية الوظيفٌة الجديدٌة له.

وبالنسبة لدفعات المورد XYZ Ltd. فقد تضمنت ٧٦٣ دفعة تتراوح بينٌ ٧٤٩.٣٤ دولار و ٣٤٦١.٩٧ دولار وان هنالك دفعة واحدة فقط بمبلغ ٢٣٤٨٩.٩٢ دولار وعند التحري تبينٌ بأن هذه الدفعة كانت تخص شركة أخرى وهي شركة XZY Ltd.. أما بالنسبة للدفعة الخاصة بالمورد ZZZ Co. فقد كانت نتيجٌة خطأ في الإدخال، إذ أن المبلغ الحقيقٌ للدفعة كان ٢٩٩٣.٤٥ دولار والذي تٌلائم مع مدى الدفعات الذي تٌراوح بينٌ ٤٤٢.٢٨ دولار و ٤٢٨١.٠٣ دولار. وبذلك مٌكن القول أن تقنيةٌ تحليلٌ النسب تعمل على تحديدٌ الشواذ الموجودة في البيانٌات والتي من المحتمل أن تمثل نشاط احتيالٌ، وبما إن الاحتيالٌ تٌضمن أنشطة مختلفة فان ذلك عٌن بأنه لكل تقنيةٌ من التقنيٌات أعلاه دور مهم في اكتشاف أنواع معينٌة من الأنشطة الاحتيٌالية.

Source: Coderre, 2002, 16 (نقلا عن نصيف الجبوري مجلة عدد ١٩ عدد٧٠)

المحاولات العراقية لتطبيق تقنية تنقيب البيانات في البيئة العراقية:

Iraqi attempts to apply data mining in the Iraqi Environment Technology

تم اجراء عملية تنقيب البيانات من قبل (زكريا واخرون – ٢٠٠٧) على بيانات القروض للمصرف الصناعي العراقي لغرض مساعدة الإدارة على اتخاذ القرار الذي يمكنها من الحكم على الزبون فيما اذا كان التعامل معه يشكل خطرا" على المصرف أم لا. وبذلك تتمكن من التركيز على الزبائن الموثوق بهم واستبعاد من هو غير موثوق به لأنه يكون السبب فيما يتحمله المصرف من خسائر ناجمة عن اقراضه بسب عدم التسديد. ومن اجل التطبيق استخدم برنامج حاسوبي يدعى خادم لغة الاستفسار المهيكلة SQL (Structured Query Language Server 2000) Server 2000 الذي يمكن تعريفه بانه نظام إدارة قاعدة البيانات Database Management System الذي تم تصميمه للمساعدة في إدارة بيانات المنظمة فهو يقدم طرائق لمسح وتحديث وادخال البيانات التي يمكن من خلالها الحفاظ على قاعدة بيانات متجددة (Gunderloy & Jorden, 2000:263)

يقدم البرنامج خوارزميتين أساسيتين لتحليل البيانات هما:

(Tiedrich, 2000:14) (Soni & Others, 2001:2) (Bloor Research, 2001:102)

١- شجرة القرار Microsoft Decision Tree (MDT)

٢- التجميع Microsoft Clustering

إن زيادة عدد المقترضين في المصرف، يجعل من الصعب على الإدارة ان تستخلص القواعد التي تحدد المخاطر الجيدة والسيئة يدويا، لذلك فان زيادة الاهتمام بتنقيب البيانات Data Mining ناتج عن كونها تهتم بالاكتشاف والتحليل عن طريق وسائل مؤتمتة أو شبة مؤتمتة. ومن تصفح اشجار القرار والتجمعات الناتجة من نماذج تنقيب البيانات من قبل المستخدم، تسمح له بإلقاء نظرة سريعة على ناتج عملية التنقيب وبالتالي يتمكن من اتخاذ قرار سريع وحكيم مستند على نماذج واضحة ومفهومة مقرونة بتفصيلات بيانية مع النسب.

الاستنتاج:

إن تقنيات التحليل المتقدم وتنقيب البيانات هي العصب الرئيسي لمعظم المؤسسات المالية بشكل عام والبنوك وشركات الاستثمار والتأمين والاتصالات بشكل خاص، وذلك نظراً لما توفره من قدرة على الاستكشاف والتنبؤ الذي يُعتبر من أهم الاحتياجات المعرفية لهذا النوع من المؤسسات، سواء من أجل تقدير وتحليل المخاطر واتخاذ ما يلزم لتجنبها أو من أجل استكشاف الفرص الاستثمارية واقتناصها قبل فوات الأوان، إضافة لما توفره من قدرة معرفية تساهم في إدارة هذه المؤسسات بطرق حديثة تعتمد بشكل أساسي على المعرفة المستكشفة والتي يتم التنبؤ بها من خلال دراسة وتحليل وتنقيب ما يتوفر من بيانات تاريخية في مجال العمل أو حتى في مجالات أخرى، ففي شركات التأمين يتم الاعتماد بشكل أساسي على دراسة وتحليل بيانات حوادث المركبات من كل الأنواع ومن ثم الاستكشاف والتنبؤ بنسبة احتمال التعرض لها من أجل احتساب أقساط التأمين المناسبة لكل نوع، وكذلك ما يحدث في نظام منح القروض في البنوك الذي يتأسس على وجود خوارزميات تصنيف وتنبؤ تقوم بفحص بيانات وسمات الشخص المتقدم للقرض وتقدير المخاطرة المترتبة على منحه إياه وفق ما يتوفر من بيانات تاريخية عن المقترضين السابقين الذين يتشابهون معه في

السمات، كالعمر ومستوى الدخل وطبيعة العمل أو المهنة وغيرها، ومدى التزامهم في السداد. وتعتبر مجالات عمل المؤسسات المالية والبنوك وشركات التأمين وحتى شركات الاتصالات من المجالات الحيوية المهمة لتقنيات تنقيب البيانات وبيئة خصبة لتطبيق العديد من تقنيات التنقيب الاستكشافية والتنبؤية التي تلزم مثل هذا النوع من المؤسسات أكثر من غيرها.

قائمة المصادر

العربية:-

١- نصيف جاسم الجبوري و صلاح هادي محمد المالي دور المحاسبة انقضائية في اكتشاف عمهيات الاحتيال العدد 70

٢- زكريا مطلك الدوري و داليا عبد الحسين احمد Data Mining 2007

الانكليزية:-

1- Avison, David & Shah, Hanifa "The Information Systems Development Life Cycle": McGraw-Hill, UK, 1997

2- Coderre, David G.; "*Variance & Ratio Analysis*", Frequencies, The Journal of Size Law Applications, Volume 1, No. 2, January 2002

3- Golden, Thomas W.; Skalak, Steven L.; Clayton, Mona M.; "*A Guide to Forensic Accounting Investigation*"; John Wiley & Sons, Inc. , First Edition , 2006.

4- Goldmann, Peter ; Kaufman, Hilton; "*Anti-Fraud Risk and Control Workbook*"; John Wiley and Sons; First Edition, 2009.

5- Gunderloy, Mike & Jorden, Joseph L. "Mastering SQL Server 2000" SYBEX Inc., U.S.A, 2000.

6- Mehta, G.S. ; Mathur, Tarun; "*Preventing Financial Fraud Through Forensic Accounting*'; Journal of the Institute of Chartered Accountants of India; Volume 55, No. 10, April 2007.

7- Rambaldi, Giacomo & Bautista, Mike: Monitoring and Evaluation: Beyond Record-Keeping "Special Reports – Suhay" July-September, 2000

8- Two Crows Corporation: Data Mining Applications, 2002. http://www.Towcrows.com

9- Services, Datapro Information Services, Gartner Group, Inc., November 2000.

10- Soni, Sanjay & Tang, Zhohui & Yang, Jim: Performance Study of Microsoft Data Mining Algorithms, Microsoft Corp., 2001.

11- Bloor Research: Databases on evaluation & comparison, 2001.

12- Wu, Jonathan: What is Data Mining?, "DM Review Magazine", EC Media Group, August 2000.

13- Tiedrech, Alan H.: Microsoft Corp. SQL Server 2000 Analysis