مجلة جامعة دمشق للعلوم الاقتصادية والسياسية

اسم المقال: دور استخدام أساليب التنقيب في البيانات في تعزيز فاعلية اكتشاف غش الإدارة "دراسة تطبيقية باستخدام خوارزمية سرب الجسيمات"
اسم الكاتب: سامر شوقل، أ.د. حسين دحدوح، أ.د. راكان رزوق
رابط ثابت: https://political-encyclopedia.org/library/1809
تاريخ الاسترداد: 2025/05/10 20:44 +03

الموسوعة السياسيّة هي مبادرة أكاديمية غير هادفة للربح، تساعد الباحثين والطلاب على الوصول واستخدام وبناء مجموعات أوسع من المحتوى العلمي العربي في مجال علم السياسة واستخدامها في الأرشيف الرقمي الموثوق به لإغناء المحتوى العربي على الإنترنت.
لمزيد من المعلومات حول الموسوعة السياسيّة - Encyclopedia Political، يرجى التواصل على info@political-encyclopedia.org

استخدامكم لأرشيف مكتبة الموسوعة السياسيّة - Encyclopedia Political يعني موافقتك على شروط وأحكام الاستخدام المتاحة على الموقع https://political-encyclopedia.org/terms-of-use

تم الحصول على هذا المقال من موقع مجلة جامعة دمشق للعلوم الاقتصادية والسياسية ورفده في مكتبة الموسوعة السياسيّة مستوفياً شروط حقوق الملكية الفكريّة ومتطلبات رخصة المشاع الإبداعي التي ينضوي المقال تحتها.

مجلة جامعة دمشق للعلوم الاقتصادية والسياسية
المجلد 39- 2- 2023: 141-161

Damascus University Journal
For The Economic And Political Sciences
Vol 39 -2-2023: 141-161

دور استخدام أساليب التنقيب في البيانات في تعزيز فاعلية اكتشاف غش الإدارة
"دراسة تطبيقية باستخدام خوارزمية سرب الجسيمات"

سامر شوقل[1]، أ. د. حسين دحدوح[2]، أ. د. راكان رزوق[3]

[1]طالب دكتوراه، قسم المحاسبة- كلية الاقتصاد- جامعة دمشق.

[2]أستاذ في قسم المحاسبة – كلية الاقتصاد – جامعة دمشق.

[3]أستاذ في قسم هندسة البرمجيات ونظم المعلومات – كلية الهندسة المعلوماتية – جامعة دمشق.

الملخص

هدف البحث إلى اختبار دور استخدام أساليب التنقيب في البيانات في تعزيز فاعلية اكتشاف الغش في القوائم المالية في بيئة الأعمال السورية، وذلك بالتطبيق على الشركات الصناعية والزراعية المدرجة في سوق دمشق للأوراق المالية والبالغ عددها (3) شركات مدرجة وذلك لتسع فترات مالية؛ من عام (2012) وحتى عام (2020) بعدد إجمالي (27) مشاهدة. وللوصول إلى هدف البحث تم تطبيق نموذج Beneish M Score كأسلوب تقليدي لاكتشاف الغش وتصنيف الشركات محل الدراسة إلى شركات محرفة (تحتوي على غش) وشركات غير محرفة (لا تحتوي على غش). ثم مقارنة النتائج مع تطبيق إحدى أساليب التنقيب في البيانات باستخدام خوارزمية سرب الجسيمات (PSO) مدعومةً بخوارزمية آلات متجهات الدعم (SVM).

وقد توصل البحث إلى مجموعة من النتائج أهمها: وجود اختلاف في النتائج بين استخدام المدقق لأساليب التدقيق التقليدية في اكتشاف الغش وبين استخدامه لأساليب التنقيب في البيانات بنسبة اختلاف قدرها (83.33%) في ضوء مجتمع الدراسة. كما أظهر البحث تفوّق أساليب التنقيب في البيانات (خوارزمية سرب الجسيمات PSO مع خوارزمية آلات متجهات الدعم SVM) في التنبؤ بغش الإدارة على النماذج الاحصائية التقليدية التي تُستخدم للتنبؤ بحالات الغش.

الكلمات المفتاحية: التنقيب في البيانات، خوارزمية سرب الجسيمات، اكتشاف الغش، نموذج Beneish M Score.

تاريخ الايداع: 2021/6/29
تاريخ القبول: 2021/9/19

حقوق النشر: جامعة دمشق سورية، يحتفظ المؤلفون بحقوق النشر بموجب **CC BY-NC-SA**

The role of using data mining methods in enhancing the effectiveness of management fraud detection "An applied study using the particle swarm algorithm"

Samer Shawqal[1]
Prof. Housin Dahdouh[2], Prof. Rakan Razouk[3]

[1]PhD Student, Accounting Department, Faculty of Economics, Damascus University.
[2]Professor Department of Accounting, Faculty of Economics, Damascus University.
[3]Professor Department of Software Engineering and Information Systems, Faculty of Informatics Engineering, Damascus University

Received: 29/6/2021
Accepted: 19/9/2021

Copyright: Damascus University- Syria,
The authors retain the copyright under a **CC BY- NC-SA**

Abstract

The research aimed to test the role of using data mining methods in enhancing the effectiveness of detecting fraud in financial statements in the Syrian business environment, by applying it to the industrial and agricultural companies listed on the Damascus Securities Exchange, which are (3) listed companies for nine financial periods; From (2012) to (2020) with a total number of (27) observations. To reach the aim of the research, the Beneish M Score model was applied as a traditional method for detecting fraud and classifying the companies under study into fraudulent companies (containing fraud) and non-fraudulent companies (containing fraud). Then, the results were compared with the result of applying data mining method using the Particle Swarm Optimization (PSO) Algorithm supported by the Support Vector machines (SVM) algorithm.The research found a set of results, The most important of them are: There is a difference in the results between the auditor's use of traditional auditing methods in detecting fraud and his use of data mining methods with a difference of (83.33%) in the light of the study sample. The research also showed the superiority of data mining methods (Particle Swarm Optimization (PSO) Algorithm with the Support Vector machines (SVM) algorithm) in predicting management fraud over traditional statistical models that are used to predict fraud cases.

Keywords: Data Mining, Particle Swarm Algorithm, Fraud Detection, Beneish M Score Model.

المقدمة:

البيانات هي نفط، والبعض يقول هي ذهب القرن الواحد والعشرين. هذا هو المصطلح الذي يُستخدم للدلالة على أهمية البيانات في العصر الحالي. وكما في حالة النفط الخام، لا يمكن استعماله والاستفادة منه إلا في حال تكريره. كذلك هي البيانات، لا يمكن الاستفادة منها إلا في حال تحليلها واستخراج ما ينفع منها ويفيد (الخياط، 2018). وفي عالم الأعمال لا تستمد تلك البيانات (ماليةً أو غير مالية) قيمتها وأهميتها إلا من خلال تدقيق الحسابات، حيث أضحى التدقيق هو الملاذ الآمن للإدارة والأطراف المستفيدة والدليل القوي على حيادة الأرقام الموجودة في القوائم المالية، وخاصةً بعد ما شهدته السنوات الأخيرة من ازدياد التلاعب في القوائم المالية، وما ترتب عليه من آثار سلبية على الاقتصاد العالمي، وإفلاس العديد من الشركات أو تدهور سمعتها بسبب تلاعب الإدارة في القوائم المالية، مثل الفضيحة التي حدثت في عام (2015) والتي هزت سمعة ومكانة واحدة من أكبر الشركات التي تعمل في مجال تكنولوجيا المعلومات وهي شركة "توشيبا" التي ضخمت أرباحها بمقدار (1.22 مليار دولار) خلال (7) سنوات!، حيث كانت أشهر عملية غش مارستها الإدارة في تاريخ شركة "توشيبا" وتاريخ الشركات اليابانية (Edwards, 2015). لذلك ازداد اهتمام المنظمات المهنية في العديد من دول العالم للاهتمام بتحديد مسؤولية المدقق الخارجي في اكتشاف التلاعب (الغش) في القوائم المالية لا سيما في عصر التطورات السريعة والمتلاحقة في مجال تقنيات المعلومات الذي أتاح فرصاً أكبر لارتكاب الغش من قبل إدارة الشركات، وخاصةً عندما تكون تلك الإدارات في موضع يسمح لها بتجاوز إجراءات الرقابة الداخلية وبالتالي التلاعب في العمليات المحاسبية بشكل مباشر أو غير مباشر، وبذلك أصبح المدقق الخارجي يواجه تحدياً كبيراً عند قيامه بمسؤولياته ذات الصلة باكتشاف الغش في القوائم المالية والتقرير عنه، وبالتالي جعلت برامج التدقيق النمطية والتقليدية المدققين الخارجيين أقل فاعلية في الاستجابة لمخاطر الغش، مما خيب آمال المستفيدين من القوائم المالية بعد تلك الانهيارات التي حصلت في العديد من اقتصاديات الدول العالمية؛ لقصور الأساليب التقليدية في اكتشاف التلاعب الذي مارسته إدارة الشركات محل التدقيق وهذا ما أكدت عليه العديد من الدراسات، منها دراسة (Appelbaum & Vasarhelyi, 2017)، ودراسة (Hoffman & Zimbelman, 2009). وفي هذا السياق يسعى الباحث لأن تكون هذه الدراسة محاولة علمية وعملية لتطبيق تقنيات حديثة في تدقيق الحسابات تستند إلى أساليب التنقيب في البيانات محاولةً متواضعة للارتقاء بمهنة التدقيق ورفع من سويتها ومواكبتها للتطورات التي أُدخلت حديثاً على هذه المهنة في العديد من الدول المتقدمة ولا مكان لها في الكثير من الدول النامية.

مشكلة البحث:

يعيش مدققو الحسابات وسط بيئة معقدة ومتشعبة من عالم الأعمال، فلم تعد تقتصر خدماتهم على تدقيق منشآت صغيرة الحجم تتسم ببساطة عملياتها وتعاملاتها مقارنةً بمنشآت أخرى ضخمة تعتمد على تقنيات المعلومات، وما ينتج عنها من بيانات بكميات هائلة على مستوى كامل القطر، مما يشكل ذلك تحدياً في وجه مدققي الحسابات من حيث قصور أساليب التدقيق التقليدية والتحليلات البسيطة أمام كم هائل من البيانات والعمليات المعقدة والمتشعبة، وبات من الصعوبةِ بمكان للمدققين أن يفو بمسؤولياتهم ذات الصلة باكتشاف الغش في ظل التطور السريع والمتلاحق في مجال تقنيات المعلومات، ولا يتأتى للمدقق أن يقدم ضماناً معقولاً بأن القوائم المالية خالية من أي تحريفات جوهرية إلّا إذا كان مواكباً للتطورات المتسارعة في بيئة الأعمال ومواكباً للأساليب المتطورة والتقنيات الحديثة في مجال التدقيق والتي باتت تُطَّبق في الدول المتقدمة ولا مكان لها في الدول النامية، حيث لم يعد للأساليب والتقنيات التقليدية في مهنة التدقيق مكانٌ أمام التطورات المتلاحقة والبيانات المعقدة والمتشعبة وخاصة عند استخدام تلك الأساليب والاجراءات التقليدية في اكتشاف التحريفات الجوهرية، سواءٌ الناتجة عن خطأ أم غش، ولذا كان من

الضرورة بمكان استخدام تقنيات حديثة تعتمد على تقنيات التنقيب في البيانات باستخدام خوارزميات قد تنامى استخدامها مؤخراً في مجال تعزيز جودة التدقيق وبالتالي تعزيز فاعلية تدقيق الحسابات (Hooda, Bawa, & Rana, 2018)، لا سيما في اكتشاف الغش في القوائم المالية الذي تمارسه إدارة الشركة محل التدقيق. وبناءً على العرض السابق يمكن صياغة مشكلة البحث في الإجابة على التساؤلات التالية:

1- هل يوجد اختلاف بين استخدام المدقق لأساليب التدقيق التقليدية وأساليب التنقيب في البيانات في اكتشاف الغش في القوائم المالية؟

2- هل يؤدي استخدام المدقق لأساليب التنقيب في البيانات إلى زيادة فاعلية التدقيق من حيث اكتشاف الغش في القوائم المالية ؟

أهداف البحث:

يهدف هذا البحث إلى اختبار دور استخدام أساليب التنقيب في البيانات في تعزيز فاعلية اكتشاف الغش في القوائم المالية، ويتم ذلك من خلال تحقيق الأهداف التالية:

1- بيان الاختلاف بين استخدام المدقق للأساليب التقليدية وأساليب التنقيب في البيانات في تعزيز فاعلية اكتشاف الغش في القوائم المالية.

2- بيان دور استخدام المدقق لأساليب التنقيب في البيانات في زيادة فاعلية اكتشاف الغش في القوائم المالية.

أهمية البحث:

انطلاقاً من إدراكنا أهمية استخدام أساليب التنقيب في البيانات في تعزيز فاعلية اكتشاف الغش في القوائم المالية يمكن تقسيم أهمية البحث إلى:

الأهمية العلمية: تكمن في أهمية تطرق البحث إلى مجال مهم وحديث في مجال تدقيق الحسابات لا سيما في أساليب اكتشاف الغش في القوائم المالية في ظل بيئة تتسم بالتطور التقني الكبير في مجال تكنولوجيا المعلومات وذلك من خلال الاستناد إلى أساليب تُستخدم في مجال هندسة المعلومات بشكل عام وأساليب التنقيب في البيانات، باستخدام خوارزميات معينة بشكل خاص. كما تنبع أهميتها العلمية بأنها تحاول التوصل إلى نتائج تفيد في إثراء الجهود البحثية التي تُجرى حالياً حول استخدام أساليب التنقيب في البيانات في مجال تدقيق الحسابات.

الأهمية العملية: يستمد هذا البحث أهميته العملية؛ لعَدّهِ محركاً مهماً ومشجعاً لاستخدام تقنيات التنقيب في البيانات من قبل مكاتب وشركات التدقيق العاملة في سورية بدلاً عن استخدامهم الأساليب التقليدية عند أداء عملية التدقيق الخارجي بشكل عام، وعند القيام بالإجراءات التي لها علاقة باكتشاف الغش أو المخالفات والتصرفات غير المشروعة بشكل خاص، فضلاً عن أهمية هذا البحث في تعزيز فاعلية التدقيق والارتقاء بالمهنة ومواجهة التحديات والتطورات المتسارعة في بيئة الأعمال مما يشكل حمايةً للمجتمع المالي، وللمدقق على حدٍ سواء.

فرضيات البحث:

انطلاقاً من مشكلة هذا البحث وتحقيقاً لأهدافه، يمكن صياغة الفرضيات التالية:

1- لا يوجد اختلاف بين استخدام المدقق لأساليب التدقيق التقليدية وأساليب التنقيب في البيانات في اكتشاف الغش.

2- لا يؤدي استخدام المدقق لأساليب التنقيب في البيانات إلى زيادة فاعلية اكتشاف الغش في القوائم المالية.

الدراسات السابقة:

1- دراسة (Nawaiseh & Abbod, 2020) بعنوان: تدقيق القوائم المالية باستخدام تقنية آلات متجهات الدعم، وتقنية الشبكات العصبية الصنعية، و تقنية الجار الأقرب- دراسة اختبارية في المملكة المتحدة وايرلندا.

هدفت هذه الدراسة إلى كشف قدرات تقنيات التنقيب في البيانات في بناء نماذج تنبؤية عند تدقيق القوائم المالية للشركات وذلك بالاعتماد على ثلاث تقنيات وهي: تقنية آلات متجهات الدعم (SVM)، وتقنية الشبكات العصبية الصنعية (ANN)، وتقنية الجار الأقرب (KNN)، حيث أظهرت الدراسة تفوق تقنية الشبكات العصبية الصنعية وتقنية آلات متجهات الدعم على تقنية الجار الأقرب في إعطاء نتائج أدق من حيث تصنيف الشركات السليمة مالياً. وكانت تقنية الشبكات العصبية الصنعية ذات مقدرة أفضل في تصنيف الشركات مقارنةً مع التقنيات الأخرى.

2- دراسة (Mollah & Sakib, 2020) بعنوان: اكتشاف الغش في القوائم المالية باستخدام نموذج Beneish M-score- دراسة تطبيقية على شركات الأدوية المدرجة في بورصة بنغلادش.

هدفت هذه الدراسة إلى تحديد قيمة نموذج Beneish M-score في اكتشاف التلاعب في القوائم المالية، حيث تم تطبيق هذا النموذج على شركات الأدوية المدرجة في بورصة بنغلادش والبالغ عددها (14) شركة خلال الفترة الممتدة من العام /2014/ وحتى العام /2018/. وتوصلت الدراسة إلى ممارسة شركات الأدوية أساليب التلاعب بالقوائم المالية وتضخيم الأرباح، وأظهرت الدراسة تدهوراً في قطاع صناعة الأدوية مما جعل المدققين والمنظمات المهنية أكثر حذراً في الكشف عن عمليات الغش في القوائم المالية.

3- دراسة (الجبلي، 2020) بعنوان: أثر استخدام الانحدار اللوجستي (logistic regression) كأحد أساليب التنقيب في البيانات في دعم الرأي المهني لمراجعي الحسابات.

هدفت هذه الدراسة الى معرفة أثر استخدام أسلوب الانحدار اللوجستي كأحد أساليب التنقيب في البيانات في دعم الرأي المهني لمدققي الحسابات حول مدى قدرة المنشأة على الاستمرارية و وجود تحريفات جوهرية في القوائم المالية بالتطبيق على الشركات المدرجة في بورصة الأوراق المالية المصرية. وقد توصلت الدراسة الى مجموعة من النتائج، أهمها: يؤثر استخدام الانحدار اللوجستي كأحد أساليب التنقيب في البيانات في دعم الرأي المهني لمراجع الحسابات حول وجود تحريفات جوهرية في القوائم المالية و في دعم الرأي المهني لمراجع الحسابات حول استمرارية المنشأة، كما توصلت إلى ضرورة توسيع نطاق الإجراءات التحليلية لتتضمن أساليب التنقيب في البيانات المناسبة لتحسين دقة التنبؤ بقدرة المنشأة على الاستمرارية وبمدى وجود تحريفات جوهرية في القوائم المالية.

4- دراسة (Yan, Wang, & Zhang, 2019) بعنوان: تطبيق تقنيات التنقيب في البيانات في تقييم مخاطر التدقيق.

هدفت هذه الدراسة إلى بيان مزايا تقنيات التنقيب في البيانات عند تقييم مخاطر التدقيق لا سيما في ظل التغيرات الكبيرة في بيئة الأعمال وظهور بيانات ضخمة وضعت المدقق الخارجي في صعوبات وتحديات أدائه لعملية التدقيق، حيث وضحت هذه الدراسة إجراءات تنفيذ عملية تقييم المخاطر باستخدام خوارزميات العنقدة (clustering) كإحدى تقنيات التنقيب في البيانات، وتوصلت الدراسة إلى أنَّ استخدام المدقق لخوارزميات العنقدة يوفر له الكثير من الجهد والوقت عند أداء عملية التدقيق لا سيما في الدقة في تحديد البيانات التي تثير الشكوك والتي تحتوي على تحريفات جوهرية، مما يخفض من خطر التدقيق.

5- دراسة (أحمد، 2019) بعنوان: دور استخدام أساليب التنقيب في البيانات لتحسين تقديرات مراقب الحسابات في مدى وجود أخطاء جوهرية بالقوائم المالية- دراسة ميدانية في بيئة الأعمال المصرية

هدفت هذه الدراسة إلى التعرف على مدى تأثير استخدام أساليب التنقيب في البيانات في تحسين تقديرات مراقب الحسابات عند قيامه بمراجعة القوائم المالية، وأجريت الدراسة من خلال استبيان وزع على عينة من مراقبي الحسابات في مصر، وقد توصلت الدراسة الى مجموعة من النتائج أهمها أنَّ استخدام تطبيقات كل من تقنية الشبكات العصبية الصنعية وتقنية الانحدار اللوجستي تدعم الرأي المهني لمراقب الحسابات حول تقييم فرض الاستمرارية، وتوصلت إلى ضرورة توسيع نطاق الإجراءات التحليلية لتتضمن أساليب التنقيب في البيانات من أجل تحسين دقة التنبؤ بمدى وجود أخطاء جوهرية في القوائم المالية.

ما يميز البحث عن الدراسات السابقة بأنه تناول دراسة دور استخدام تقنيات التنقيب في البيانات في فاعلية التدقيق في البيئة السورية من خلال بيانات فعلية وليس من خلال استبيان وُزّع على عينة من المدققين كما فعلت دراسة (أحمد، 2019). كما اختلفت متغيرات هذا البحث عن متغيرات الدراسات السابقة (كما هو موضح في فقرة الأنموذج)؛ حيث ركز هذا البحث على فاعلية التدقيق (كمتغير تابع) من خلال الاهتمام بالغش المتعلق بالإدارة والذي يعتبر أكثر خطورة من بقية التحريفات المتعمدة في القوائم المالية وبالتالي أكثر خطورة من بقية التحريفات الجوهرية ككل؛ لأن الإدارة عادةً ما تكون في موضع يُمكنها من التلاعب بشكل مباشر أو غير مباشر في السجلات المحاسبية أو تجاوزها لإجراءات الرقابة المصممة لمنع عمليات الغش المشابهة من قبل موظفين آخرين، وكذلك اختلف البحث في الأسلوب المستخدم في التنقيب في البيانات (كمتغير مستقل) حيث استخدم البحث خوارزمية سرب الجسيمات (PSO) مدعومةً بخوارزمية (SVM) في حين أنَّ الدراسات السابقة استخدمت أساليب أخرى.

منهجية البحث:

في سبيل تحقيق أهداف البحث واختبار فرضياته، قام الباحث باتباع ما يلي:

أولاً: المنهج الوصفي: لتطوير الإطار النظري للدراسة من خلال دراسة وتحليل الاصدارات المهنية والدراسات السابقة الحديثة في هذه المجال لاشتقاق وصياغة فروض البحث.

ثانياً: المنهج التحليلي: لاختبار فروض البحث من خلال الدراسة التطبيقية والتحليلية، وذلك باتباع الخطوات التالية:

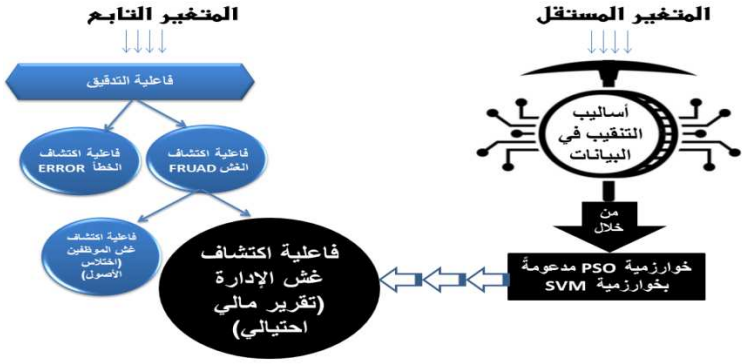
1- استخدام أحد النماذج الاحصائية كأحد الأساليب التقليدية التي يستخدمها المدقق الخارجي للتصنيف المبدئي للشركات الصناعية السورية المدرجة في سوق دمشق للأوراق المالية محل الدراسة إلى قوائم تحتوي على غش وقوائم لا تحتوي على غش.

2- استخدام أحد أساليب التنقيب في البيانات على نفس الشركات في الخطوة السابقة، حيث تمثل الشركات المصنفة في الخطوة السابقة مدخلات برنامج التنقيب في البيانات وسيقوم البرنامج بالتعلم من البيانات المدخلة وبناء نموذج التصنيف المتعلق بالغش في القوائم المالية (مرحلة تدريب الخوارزمية).

3- اختبار مدى صحة نموذج التنقيب في البيانات باستخدام فترات مالية أخرى بخلاف التي أُستخدمت في الخطوة الأولى (مرحلة اختبار الخوارزمية).

4- مقارنة النتائج التي ظهرت في الخطوة رقم (3) والتي تم فيها استخدام أسلوب التنقيب في البيانات مع النتائج التي ظهرت في الخطوة رقم (1) والتي تم فيها استخدام النموذج الاحصائي التقليدي.

أنموذج البحث*:



إنَّ هذا الأنموذج هو من إعداد الباحث وذلك بالاستناد إلى عدد من الأبحاث والدراسات السابقة وإلى معيار التدقيق الدولي رقم (240) المتعلق بـ "مسؤوليات المدقق ذات العلاقة بالغش عند تدقيق القوائم المالية".

❖ الإطار النظري

أولاً: مفهوم التنقيب في البيانات:

يشير مصطلح التنقيب في البيانات إلى استخراج المعرفة (Knowledge discovery) المخبأة ضمن كميات كبيرة من البيانات (رزوق، 2013، ص3)، كما يشير أيضاً إلى عملية البحث داخل كميات كبيرة من البيانات للكشف عن العلاقات التي لم يتم كشفها سابقاً بين عناصر البيانات (Dawalekar, Gharat, Gaikwad, & Kundale, 2019)، ويشير أيضاً إلى النشاط الذي يقوم باستخراج المعلومات المتواجدة في كميات كبيرة من البيانات، بهدف البحث عن قواعد معرفية (Knowledge rules) واكتشاف الحقائق الخفية الواردة في قواعد البيانات (Nageswari, Goel, 2019, p.1). وقد أشار البعض أنَّ مصطلح التنقيب في البيانات هو مصطلح غير دقيق؛ فلو تم إجراء مقارنة مع عملية التنقيب عن معدن معين، كالذهب مثلاً، فإنَّ هذا المعدن يكون موجود أصلاً ضمن الرمال أو الصخور التي تجرى معالجتها لعزل المعدن عن الرمل أو الصخر، في حين أنَّ التنقيب في البيانات يقوم باستخراج معلومات غير موجودة بشكل واضح ضمن كمية البيانات الضخمة التي يجري البحث ضمنها، ومن ثم فالعملية ليست عزل بعض المكونات "المفيدة" عن غيرها من المكونات "غير المفيدة"، بل إنها عملية استخراج معارف جديدة اعتماداً على البيانات، وبالتالي يكون مصطلح "اكتشاف المعرفة" هو المصطلح الأكثر تعبيراً، لكن يتم استخدام مصطلح التنقيب في البيانات لأنه المصطلح الشائع (رزوق، 2013، ص1).

ثانياً: الهدف من التنقيب في البيانات:

أشارت الدراسات الأولية مثل دراسة (MCluskey & Anand, 1999, p.218) إلى أنَّ أهمية التنقيب في البيانات ترجع إلى تحقيق مجموعة من الأهداف والتي تتمثل في التصدي للعديد من التحديات التي تفرضها البيانات الكبيرة، والمتباينة، والتاريخية، فضلاً عن تلبية احتياجات المستخدمين من تجهيز وتهيئة للبيانات، ومشاركة خبراء المجالات المختلفة ومعارفهم في عملية التنقيب.

وقد لخصت دراسة (Silwattananusarn & Tuamsuk, 2012, p.16) أهداف التنقيب في البيانات في هدفين رئيسيين: الهدف الأول هو "التنبؤ"؛ والذي يعتمد على استخدام بعض المتغيرات في مجموعات البيانات من أجل التنبؤ بقيم غير معروفة للمتغيرات

الأخرى ذات الصلة. أما الهدف الثاني فهو "الوصف"؛ وينطوي على إيجاد اتجاهات (trends) وأنماط (patterns) في البيانات يمكن للإنسان فهمها مثل الكشف عن القيم الشاذة وقواعد الارتباط في قواعد البيانات الضخمة.

ثالثاً: أهم الخوارزميات المستخدمة في التنقيب في البيانات:

فيما يلي من أهم الخوارزميات التي تم الاعتماد عليها بشكل خاص في مجالات المحاسبة والتدقيق:- خوارزمية الجار الأقرب K-Nearest Neighbor: تعتمد هذه الخوارزمية على مدخل التصنيف القائم على الحالات المتشابهة بالاعتماد على مقياس التشابه Similarity measure(مثنى ورياض، 2012، ص147).

- **خوارزمية آلات متجهات الدعم Support Vector Machine:** وهي أداة تنبؤيه للتصنيف والانحدار المبنية على استخدام نظرية تعلم الآلة لتعظيم الدقة التنبؤية وفي نفس الوقت تلافي الأخطاء العشوائية للبيانات Over-fit to the data (Jakkula, 2006, p.1).

- **خوارزمية أشجار القرار Decision Trees:** تعد شجرة القرارات من أهم الخوارزميات التي تستخدم في تصنيف البيانات وهي أيضاً من أدوات دعم اتخاذ القرار وهي تعتمد على تكوين خريطة تبدأ بالملاحظات وتنتهي بالنتائج المحتملة (Shama & Panigrahi, 2012, p.40)

- **خوارزمية الغابات العشوائية Random Forests:** يعتبر أول من اقترح خوارزمية الغابات العشوائية هو العالم (Breiman)، حيث تعتمد على مجموعة فرق من أشجار القرار المستخدمة في عمليات التصنيف والانحدار (Breiman, 2001).

- **خوارزمية سرب الجسيمات Particle Swarm Optimization:** تم تطوير هذه الخوارزمية من قبل الدكتور Eberhart (مهندس الكترون) وKennedy (عالم نفسي اجتماعي) في عام 1995، وتستند هذه الخوارزمية على اثنين من العلوم الأساسية وهما العلوم الاجتماعية وعلوم الحاسب، حيث تعتمد PSO على مفهوم ذكاء السرب حيث أثبت العلماء وجود أنواع متعددة من أشكال الذكاء الجماعي منبثقة من مجتمع الأسماك والطيور والحشرات، وأن هذا الذكاء الجماعي يعد ملهماً لنظام جديد في علوم الحاسوب وذكاء السرب (Del, Venayagamoorthy, Mohagheghi, &Harley, 2008, p.172). وقد تنامى استخدام خوارزمية سرب الجسيمات في التطبيقات العملية في شتى المجالات، حيث ذكرت دراسة (Poli &Blackwell, 2007) أكثر من (26) مجالاً علمياً يتم فيها استخدام خوارزمية سرب الجسيمات، يذكر الباحث ما يخص منها مجالات تحليل البيانات والمحاسبة والتدقيق والاقتصاد؛ كاستخدامها في تطبيقات الرقابة (controlling)، وتطبيقات العنقدة (clustering)، والتصنيف(classification)، وتنقيب البيانات (Data mining)، والشبكات العصبية، والتنبؤ(predictions)، والنمذجة (modeling)، والعرض (visualization)، والتمويل (finance) والاقتصاد (economics). فضلاً عن إمكانية استخدامها في تطبيقات تتطلب متطلبات خاصة كتطبيقات أمنية وعسكرية، وطبية والكترونية.

رابعاً: مفهوم الغش ومسؤولية المدقق عن اكتشافه:

أشار معيار التدقيق الدولي رقم /240/ (IFAC, 2020, P.174) إلى أنَّ الغش هو فعل متعمد يقوم به فرد أو عدة أفراد في الإدارة أو أولئك المكلفون بالحوكمة أو الموظفون أو أطراف خارجية، ويتعلق هذا الفعل باستخدام الخداع من أجل الحصول على مصلحة غير عادلة أو غير قانونية ينتج منه تحريف في القوائم المالية. وبالرغم من أنَّ الغش مفهوم قانوني واسع، فإن المدقق مهتم بالأفعال والتصرفات الاحتيالية والتي تسبب تحريفات هامة في التقارير المالية، ويمكن أن يُنظر إلى الغش على أنه تلاعب. ويشمل التلاعب مفهومين أساسيين؛ الأول: يُقصد به التلاعب بالحسابات بهدف تحقيق ربح صوري وإظهار أداء المشروع على أنه جيد، وذلك لتضليل وخداع مستخدمي القوائم المالية من مساهمين ومقرضين ومستثمرين وغيرهم. والثاني: مجموعة الأفعال التي يقوم بها المختلس لتغطية واقعة الاختلاس بحيث تظهر بصورة صحيحة، ومن أمثلته تسجيل عمليات وهمية وتزوير السجلات

والوثائق والإضافة والحذف والتزوير وغير ذلك (القاضي، ودحدوح، وقريط، 2013، ص224). وقد ميز معيار التدقيق الدولي رقم /240/ بين نوعين من التحريفات الناتجة عن الغش وهما تحريفات جوهرية ناتجة عن إعداد قوائم مالية مضللة والتي تتم عن طريق الإدارة وتحريفات جوهرية ناتجة عن سوء استخدام (اختلاس) الأصول والتي تتم عن طريق الموظفين. كما أشار المعيار إلى أنَّ الخطر الناتج من عدم قدرة المدقق على اكتشاف الغش المرتكب من جانب الإدارة أكبر من الخطر الناتج من عدم قدرته على اكتشاف الغش المرتكب من جانب الموظفين. ويرجع سبب ذلك إلى أنَّ الإدارة بموقعها الوظيفي تستطيع التلاعب المباشر وغير المباشر في السجلات المحاسبية، وتقديم معلومات مالية محرفة، وتجاوز الإجراءات الرقابية. ولهذا السبب ركز الباحث في هذا البحث على هذا النوع من الغش (غش الإدارة).

أما من ناحية مسؤولية المدقق عن اكتشاف الغش، فقد قامت المنظمات المهنية بإصدار العديد من المعايير خلال السنوات الأخيرة لمحاولة توضيح مسؤولية المدقق عن اكتشاف الغش، فعلى مستوى المعهد الامريكي للمحاسبين القانونيين فقد تمثل جهوده بإصدار العديد من النشرات التي تناولت مسؤولية المدقق عن اكتشاف الغش في القوائم المالية، كان آخر هذه النشرات هو المعيار الامريكي رقم /SAS No.99/ الذي أُصدر عام (2002) والذي أصبح المدقق بموجبه مسؤولاً عن تخطيط وأداء عملية التدقيق للحصول على تأكيد معقول فيما ما إذا كانت القوائم المالية خالية من الأخطاء الجوهرية سواءً الناتجة عن خطأ أم عن غش (AICPA, 2002). أما على مستوى الاتحاد الدولي للمحاسبين، فقد كانت آخر جهوده في هذا المجال، إصدار معيار التدقيق الدولي رقم (240)، والذي أوضح أنَّ المسؤولية الأساسية لاكتشاف ومنع الغش تقع على عاتق إدارة المنشأة والمسؤولين عن الحوكمة، أما مسؤولية المدقق فتتركز حول توفير تأكيد معقول (Reasonable Assurance) بأنَّ القوائم المالية ككل خالية من أية تحريفات جوهرية سواءً الناتجة عن خطأ أو الناتجة عن غش.

❖ الدراسة التطبيقية

أولاً- توصيف متغيرات البحث:

يمكن توصيف متغيرات البحث على النحو التالي:

1- المتغير المستقل: ويتمثل في استخدام المدقق الخارجي لأساليب التنقيب في البيانات من خلال تطبيق خوارزمية سرب الجسيمات (PSO) مدعومةً بخوارزمية آلات متجهات الدعم (SVM) على مجتمع الدراسة.

2- المتغير التابع: ويتمثل في تقييم فاعلية اكتشاف الغش في القوائم المالية من قبل المدقق الخارجي. ويتم قياسه من خلال مقارنة النتائج باستخدام إحدى أساليب التدقيق التقليدية (باستخدام أسلوب Beneish M score) مع النتائج باستخدام إحدى أساليب التنقيب في البيانات (خوارزمية SVM-PSO).

ثانياً- مجتمع البحث:

يتكون مجتمع البحث من الشركات الصناعية والزراعية المدرجة في سوق دمشق للأوراق المالية، ويُعزى اقتصار الدراسة فقط على الشركات المذكورة المدرجة (دون الشركات غير المدرجة)؛ لأنَّ الشركات المدرجة تكون خاضعةً لرقابة وإشراف هيئة الأوراق والأسواق المالية ويتم انتقائها وفق اعتبارات وشروط خاصة قد تكون صارمة لا تؤهل الشركات الأخرى للإدراج في سوق دمشق للأوراق المالية، مما سيعطي مصداقية ودقة أكبر للنتائج التي سيتم التوصل لها.

فاستخدم الباحث في هذا البحث المجتمع المذكور والذي يتألف من جميع الشركات الصناعية والزراعية المدرجة في سوق دمشق للأواق المالية والبالغ عددها (3) شركات[1] وذلك لتسع فترات مالية، من عام (2012) وحتى عام (2020) بعدد إجمالي (27) مشاهدة. وقد تم تقسيم المجتمع إلى مجموعتين؛ تتمثل المجموعة الأولى الفترات المستخدمة في مرحلة التدريب والتعلم (Training data) لخوارزمية التنقيب في البيانات في الفترة الممتدة بين عامي (2012) و(2018). أما المجموعة الثانية؛ فتتمثل في الفترات المستخدمة في مرحلة الاختبار (Test data) لخوارزمية التنقيب في البيانات في الفترة الممتدة بين عامي (2019) و(2020).

ثالثاً– مصادر جمع البيانات:

قام الباحث بالحصول على البيانات اللازمة لقياس المتغير التابع من خلال المعلومات الواردة بالقوائم المالية للشركات الصناعية والزراعية المدرجة في سوق دمشق للأوراق المالية والمرفق بها الإيضاحات المتممة لها، وفي سبيل الحصول على تلك القوائم اعتمد الباحث على المواقع الالكترونية التالية و التي تهتم بتوفير المعلومات المالية عن مجتمع الدراسة: موقع سوق دمشق للأوراق المالية (www.dse.gov.sy) وموقع هيئة الأوراق والأسواق المالية (www.scfms.sy). ثم قام الباحث بإعادة تصنيف القوائم المالية من تحليل ودمج وتفصيل حساباتها بما يخدم أغراض استخراج المؤشرات اللازمة لقياس المتغير التابع.

رابعاً– خطوات البحث التطبيقية:

1– استخدام نموذج إحصائي للتصنيف المبدئي لشركات عينة التدريب:

قام الباحث هنا باستخدام نموذج إحصائي للتصنيف المبدئي للقوائم المالية لعينة التدريب (Training sample) للشركات الصناعية والزراعية المدرجة في سوق دمشق للأوراق المالية، حيث استخدمَ الباحث نموذج Beneish M-score لكونه من أكثر النماذج شيوعاً في التنبؤ بالغش (fraud)؛ وهو نموذج رياضي يستخدم ثماني نسب مالية لاكتشاف فيما إذا كانت القوائم المالية للمنشأة محرفة أم لا من قبل إدارة الشركة محل التدقيق. وهو يتشابه مع نموذج Altman Z-score ولكنه يركز على اكتشاف الغش في القوائم المالية في حين أن نموذج Altman Z-score يركز على التنبؤ بالعسر المالي للشركة، وهذا ما أكدت عليه العديد من الدراسات؛ كدراسة (Mollah *et al*., 2020)، ودراسة (MacCarthy, 2017)، ودراسة (Repousis, 2016).

من جهة أخرى اتفقت هذه الدراسات في الإشارة بأن نموذج Beneish M Score يستخدم ثمانية متغيرات تمثل مؤشرات (index) تعكس الغش في القوائم المالية، حيث تزيد احتمالات التلاعب في القوائم المالية عندما تعكس تلك القوائم متغيرات جوهرية في حسابات العملاء، وهامش الربح، وانخفاض مؤشر الأصول، ونمو المبيعات، وزيادة الاستحقاقات، وحسابات المصاريف، وحسابات الديون. ويوضح هذا النموذج مقارنة كل مؤشر على حدى، ومن ثم حساب النتيجة النهائية (M-score) وتحديد تصنيف القوائم المالية على أنها تحتوي غش أم لا. ويأخذ هذا النموذج الصيغة التالية:

$$\text{M-score}=-4.84+0.92\text{DSRI}+0.528\text{GMI}+0.404\text{AQI}+0.892\text{SGI}+0.115\text{DEPI}-0.172\text{SGAI}+4.679\ \text{TATA}-0.327\ \text{LVGI}$$

حيث أنَّ:

1) DSRI: Days Sales in Receivables Index مؤشر معدل دوران العملاء
2) GMI: Gross Margin Index مؤشر هامش الربح
3) AQI: Asset Quality Index مؤشر جودة الأصول

[1] http://www.dse.gov.sy/issuer-company/index

4) SGI: Sales Growth Index مؤشر نمو المبيعات

5) DEPI: Depreciation Index مؤشر الاهتلاك

6) SGAI: Sales, General and Administrative Expenses Index مؤشر المصروفات البيعية والإدارية والتسويقية

7) TATA: Total Acrruals to Total Assets Index مؤشر الاستحقاقات

8) LVGI: Leverage Index مؤشر الرفع المالي

ويتم قياس كل متغير من المتغيرات السابقة على النحو التالي:

1- مؤشر معدل دوران العملاء (DSRI):

يقيس هذا المؤشر النسبة بين حسابات العملاء إلى المبيعات ومقارنة هذه النسبة بالسنة السابقة، وإن ارتفاع هذا المؤشر يشير إلى تضخيم الأرباح وبالتالي التلاعب بها (Mollah *et al*, 2020, p.511).

$$\text{مؤشر معدل دوران العملاء} = \frac{\text{حسابات عملاء السنة الحالية} \div \text{مبيعات السنة الحالية}}{\text{حسابات عملاء السنة السابقة} \div \text{مبيعات السنة السابقة}}$$

2- مؤشر هامش الربح (GMI):

يقيس هذا المؤشر النسبة بين هامش الربح إلى المبيعات للسنة السابقة على هامش الربح إلى مبيعات السنة الحالية، وإن زيادة هذا المؤشر عن (1) يشير إلى تلاعب إدارة الشركة في الأرباح (Mollah *et al*, 2020, p.512).

$$\text{مؤشر هامش الربح} = \frac{\text{هامش الربح للسنة السابقة} \div \text{مبيعات السنة السابقة}}{\text{هامش الربح للسنة الحالية} \div \text{مبيعات السنة الحالية}}$$

3- مؤشر جودة الأصول (AQI):

ويحسب من خلال نسبة الأصول الثابتة بعد استبعاد الممتلكات والمصانع والمعدات (PPE) إلى إجمالي الأصول، وإن زيادة هذا المؤشر عن (1) يشير إلى تورط الشركة في سياسة تأجيل النفقات (أي رسملة النفقات) وبالتالي يعكس انخفاض جودة الأصول وارتفاع احتمال الغش (Mollah *et al*, 2020, p.512).

$$\text{مؤشر جودة الأصول} = \frac{1 - (\text{(الأصول المتداولة للسنة الحالية + الأصول الثابتة عدا ال م.م.م للسنة الحالية)} \div \text{(إجمالي الأصول للسنة الحالية)})}{1 - (\text{(الأصول المتداولة للسنة السابقة + الأصول الثابتة عدا ال م.م.م للسنة السابقة)} \div \text{(إجمالي الأصول للسنة السابقة)})}$$

4- مؤشر نمو المبيعات (SGI):

ويقيس هذا المؤشر النمو في المبيعات للسنة الحالية مقارنةً بالسنة السابقة، وإنَّ زيادة هذا المؤشر عن (1) يشير إلى احتمال أنَّ المستثمرين والمساهمين يمارسون ضغوطات على إدارة الشركة في نمو المبيعات، مما قد تلجأ الإدارة إلى التلاعب بإيراداتها كي توهم المستثمرين والمساهمين بأنَّ تحسناً طرأ على أدائها (Mollah *et al*, 2020, p.512).

$$\text{مؤشر نمو المبيعات} = \frac{\text{مبيعات السنة الحالية}}{\text{مبيعات السنة السابقة}}$$

5- مؤشر الاهتلاك (DEPI):

إنَّ زيادة هذا المؤشر عن (1) يشير إلى احتمال اهتلاك الأصول بمعدل بطيء بهدف زيادة الأرباح، وبالتالي المغالاة في تقدير العمر الإنتاجي للأصول (Maccarthy, 2017, p.162).

$$\text{مؤشر الاهتلاك} = \frac{\text{مصروف الاهتلاك للسنة السابقة} \div \text{الممتلكات والمصانع والمعدات (بالصافي للسنة السابقة)+ مصروف الاهتلاك للسنة السابقة}}{\text{مصروف الاهتلاك للسنة الحالية} \div \text{الممتلكات والمصانع والمعدات (بالصافي للسنة الحالية)+ مصروف الاهتلاك للسنة الحالية}}$$

6- مؤشر المصروفات البيعية والادارية والتسويقية (SGAI):

تقيس هذه النسبة نسبة المصروفات البيعية والادارية والتسويقية إلى المبيعات، وإنَّ زيادة هذا المؤشر عن (1) يشير إلى احتمالية غش الإدارة من خلال تأجيل المصروفات.

$$\text{مؤشر SGAI} = \frac{\text{المصروفات البيعية والادارية والتسويقية للسنة الحالية} \div \text{المبيعات للسنة الحالية}}{\text{المصروفات البيعية والادارية والتسويقية للسنة السابقة} \div \text{المبيعات للسنة السابقة}}$$

7- مؤشر الاستحقاقات (TATA):

يقيس هذه المؤشر نسبة التغير في صافي رأس المال العامل عدا النقدية مخصوماً منه مصاريف الاهتلاك والاطفاء، وإنَّ ارتفاع هذا المؤشر يشير إلى احتمالية غش الإدارة من خلال التلاعب بالاستحقاقات (Repousis, 2016, p.1066).

$$\text{مؤشر الاستحقاقات} = \frac{\Delta\text{صافي رأس المال العامل} - \Delta\text{النقدية} + \Delta\text{ضريبة الدخل المستحقة} + \Delta\text{الجزء المستحق من الديون طويلة الأجل} - \text{مصروف اهتلاك السنة الحالية}}{\text{إجمالي الأصول للسنة الحالية}}$$

8- مؤشر الرفع المالي (LVGI):

يقيس هذا المؤشر نسبة إجمالي الديون إلى إجمالي الأصول، وكلما زادت النسبة عن (1) فهذا يشير إلى غش الإدارة من خلال التلاعب في اتفاقيات الاقتراض وشراء الأصول (Mollah *et al*, 2020, p.513).

$$\text{مؤشر الرفع المالي} = \frac{\text{إجمالي الالتزامات للسنة الحالية} \div \text{إجمالي الأصول للسنة الحالية}}{\text{إجمالي الالتزامات للسنة السابقة} \div \text{إجمالي الأصول للسنة السابقة}}$$

ونتيجة لذلك، قام الباحث بتطبيق النموذج على شركات عينة الدراسة عن الفترة ما بين عامي (2012) حتى عام (2020)، وفي النهاية تم تصنيف هذه العينة من الشركات الصناعية والزراعية المدرجة إلى شركات محرفة (Fraud) وشركات غير محرفة (NoFraud). وطبقاً لنموذج Beneish M-score؛ تكون الشركة محرفة إذا كانت قيمة M-score أكبر من (−2.22) وتكون الشركة غير محرف إذا كانت القيمة أقل أو تساوي (−2.22). وبالتالي هو نموذج احتمالي (Mollah *et al*, 2020, p.513)؛ أي كلما كانت قيمة M-Score أكبر، كلما كان هناك احتمال أكبر بأنَّ إدارة الشركة تمارس الغش.

ويمكن توضيح نتائج التطبيق من خلال الجداول التالية:

الجدول رقم (1) تصنيف شركات عينة التدريب وفق نموذج Beneish لعام 2012

شركة NAMA	شركة ABC	شركة AVOC	ID\Index
1.2608	0.6260	0.4559	مؤشر DSRI
1.1459	0.9310	0.6038	مؤشر GMI
0.9425	0.9486	0.8207	مؤشر AQI
0.9479	1.3314	1.2492	مؤشر SGI
0.8911	21.2717	0.9251	مؤشر DEPI
1.2215	1.6558	0.7158	مؤشر SGAI
0.0082	0.0644	0.1707-	مؤشر TATA
1.0687	1.1299	1.3615	مؤشر LVGI
2.2674-	0.1083-	3.9166-	M-score
nofraud	fraud	nofraud	التصنيف classification

حيث أنَّ: AVOC: هي الشركة الأهلية لصناعة الزيوت النباتية، ABC: هي شركة اسمنت البادية، وNAMA: هي الشركة الهندسية الزراعية للاستثمارات نماء.

الجدول رقم (2) تصنيف شركات عينة التدريب وفق نموذج Beneish لعام 2013

شركة NAMA	شركة ABC	شركة AVOC	ID\Index
13.4428	1.2005	1.4528	مؤشر DSRI
0.3473-	1.0079	0.6980	مؤشر GMI
1.1622	0.9840	0.7564	مؤشر AQI
0.0765	0.9942	0.6417	مؤشر SGI
1.0487	0.0326	0.9243	مؤشر DEPI
80.4989	3.0098	1.2636	مؤشر SGAI
0.1381	0.0139	0.0916	مؤشر TATA
1.2140	1.0216	0.9477	مؤشر LVGI
5.5943-	2.7020-	2.2494-	M-score
nofraud	nofraud	nofraud	التصنيف classification

الجدول رقم (3) تصنيف شركات عينة التدريب وفق نموذج Beneish لعام 2014

شركة NAMA	شركة ABC	شركة AVOC	ID\Index
0.5695	0.9518	0.6298	مؤشر DSRI
0.2083-	1.8134	1.4726	مؤشر GMI
1.0444	0.9162	0.8160	مؤشر AQI
1.4393	1.0201	1.5561	مؤشر SGI
2.3110	1.5096	0.9198	مؤشر DEPI
0.0850	0.6960	1.1036	مؤشر SGAI
0.0416-	0.0044	0.0071-	مؤشر TATA
1.0881	1.3901	1.2550	مؤشر LVGI
3.0198-	2.1069-	2.2930-	M-score
nofraud	fraud	nofraud	التصنيف classification

الجدول رقم (4) تصنيف شركات عينة التدريب وفق نموذج Beneish لعام 2015

شركة NAMA	شركة ABC	شركة AVOC	ID\Index
1.3595	0.2398	0.7258	مؤشر DSRI
7.7114-	0.5223	1.3136	مؤشر GMI
1.0164	0.9019	0.7139	مؤشر AQI
0.8204	1.6576	1.3305	مؤشر SGI
1.3913	0.6497	0.8992	مؤشر DEPI
0.6687	1.1346	1.0921	مؤشر SGAI
0.0546-	0.0566	0.0308-	مؤشر TATA
1.0794	1.4870	1.0216	مؤشر LVGI
7.0818-	2.8427-	2.5661-	M-score
nofraud	nofraud	nofraud	التصنيف classification

الجدول رقم (5) تصنيف شركات عينة التدريب وفق نموذج Beneish لعام 2016

ID\Index	شركة AVOC	شركة ABC	شركة NAMA
مؤشر DSRI	1.0604	1.3780-	0.7250
مؤشر GMI	0.5845	0.8340	1.3997-
مؤشر AQI	0.6721	0.9212	0.9923
مؤشر SGI	0.9429	1.5313	1.2861
مؤشر DEPI	0.8804	0.8822	0.4035
مؤشر SGAI	0.7866	1.2066	0.5302
مؤشر TATA	0.0273	0.0681	0.0076
مؤشر LVGI	0.7712	1.3873	0.9930
M-score	2.6018-	4.1704-	3.6981-
التصنيف classification	nofraud	nofraud	nofraud

الجدول رقم (6) تصنيف شركات عينة التدريب وفق نموذج Beneish لعام 2017

ID\Index	شركة AVOC	شركة ABC	شركة NAMA
مؤشر DSRI	0.6043	1.5298-	0.5231
مؤشر GMI	2.7721	1.2150	0.5613
مؤشر AQI	0.9187	0.9344	0.9489
مؤشر SGI	1.6546	1.9252	1.8810
مؤشر DEPI	0.8890	1.8151	2.4603
مؤشر SGAI	0.6072	0.6587	0.5988
مؤشر TATA	0.5030	0.0213	0.0102-
مؤشر LVGI	1.5465	0.5034	0.9627
M-score	0.8723	3.4808-	2.1837-
التصنيف classification	fraud	nofraud	fraud

الجدول رقم (7) تصنيف شركات عينة التدريب وفق نموذج Beneish لعام 2018

ID\Index	شركة AVOC	شركة ABC	شركة NAMA
مؤشر DSRI	0.9077	1.1170	0.8052
مؤشر GMI	1.4284	0.7955	0.9123
مؤشر AQI	0.9946	0.7288	0.9499
مؤشر SGI	1.0967	1.5811	1.1567
مؤشر DEPI	0.8761	1.2568	1.1681
مؤشر SGAI	0.9473	0.5409	1.0026
مؤشر TATA	0.0784	0.1430	0.2436-
مؤشر LVGI	0.5408	0.6015	0.9354
M-score	1.7426-	1.1638-	3.6857-
التصنيف classification	fraud	fraud	nofraud

الجدول رقم (8) تصنيف شركات عينة الدراسة وفق نموذج Beneish لعام 2019

ID\Index	شركة AVOC	شركة ABC	شركة NAMA
مؤشر DSRI	1.0241	0.7612	26.6828
مؤشر GMI	0.8203	1.0394	0.9395
مؤشر AQI	0.8087	0.7438	0.9619
مؤشر SGI	0.9243	1.2747	0.0732
مؤشر DEPI	0.9323	4.9690	4.1052
مؤشر SGAI	1.0348	0.6198	3.6114
مؤشر TATA	0.1275-	0.2914	0.0362-
مؤشر LVGI	3.4987	1.0333	1.0289
M-score	4.1249-	0.6629-	20.0034
التصنيف classification	nofraud	fraud	fraud

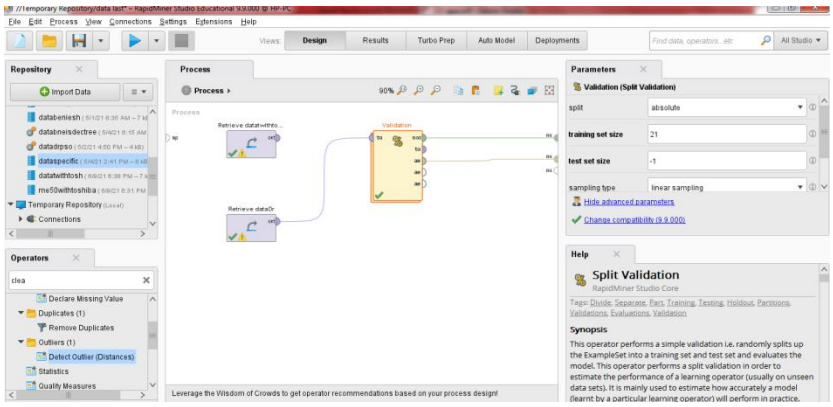
الجدول رقم (9) تصنيف شركات عينة الدراسة وفق نموذج Beneish لعام 2020

ID\Index	شركة AVOC	شركة ABC	شركة NAMA
مؤشر DSRI	0.2819	1.0783	0.0115
مؤشر GMI	0.2389	0.7304	0.9845
مؤشر AQI	0.3185	0.6884	0.8975
مؤشر SGI	3.5282	2.3231	19.4686
مؤشر DEPI	1.2579	0.8795	0.2799
مؤشر SGAI	0.5194	0.9569	0.2374
مؤشر TATA	0.5005	0.3732	0.0079-
مؤشر LVGI	1.3784	1.3004	0.8942
M-score	0.7678	0.1456	13.0808
التصنيف classification	fraud	fraud	fraud

2- تطبيق نموذج أسلوب التنقيب في البيانات باستخدام خوارزمية PSO:

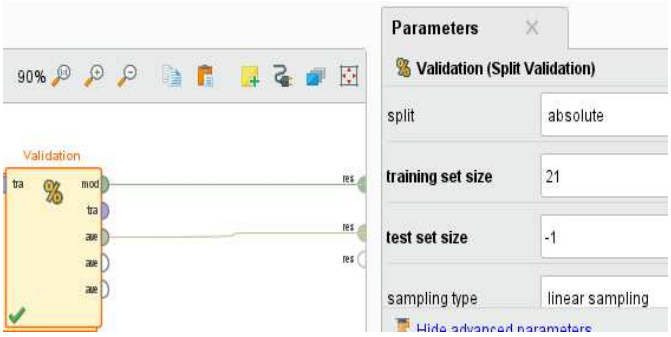
اقترح الباحث استخدام خوارزمية (PSO) مع خوارزمية (SVM)، وهو نموذج شائع الاستخدام في البرامج المتخصصة في التنقيب في البيانات وتحليلها. وتم استخدام النموذج المختلط (Hybrid) من الخوارزميات حتى يتم رفع مستوى أسلوب المعالجة لدى الخوارزمية وتعويض نقاط الضعف فيها من خلال الاستفادة من نقاط القوة في الخوارزمية الأخرى. فبعد أن تقوم خوارزمية (PSO) بتصنيف البيانات (Data set) إلى فئتين؛ محرفة وغير محرفة. تأتي خوارزمية (SVM) لتقوم بفصل هاتين الفئتين عبر شعاعين (Vectors)، محاوِلةً تعظيم الهامش (Margin) بين هذين الشعاعين لتحسين دقة التصنيف وتخفيض الوقوع في خطأ التصنيف للبيانات المتداخلة.

ويوضح الشكل رقم (1) الأداة المستخدمة لتطبيق خوارزمية PSO-Svm المؤلفة من (5) أقسام؛ قسم مصادر البيانات (Repository)، وقسم المُشغّلات (Operators) الذي يتم من خلاله اختيار أسلوب معالجة البيانات، وقسم تنفيذ العملية (Process) الذي يتم من خلاله تنفيذ العملية، وقسم إدخال المعاملات (Parameters) الذي يتم من خلاله إدخال معاملات كل مُشغل من المشغلات، وقسم المساعدة (help).



الشكل رقم (1) نافذة الأداة (RapidMiner studio 9.9)

وتم تقسيم البيانات باستخدام طريقة المعاينة الخطية (linear sampling) من خلال (validation) الأداة إلى بيانات تدريب من عام (2012) وحتى عام (2018) والبالغ عددها (21) مشاهدة كما يوضح الشكل رقم (2)، والبيانات المتبقية تمثل بيانات الاختبار وهي عامي (2019) و (2020) بعدد (6) مشاهدات.



الشكل رقم (2) تقسيم البيانات (Data Validation)

تمثل بيانات فترة التدريب (Training Data) مدخلات خوارزمية SVM-PSO، حيث قام الباحث بإعادة هندسة بيانات المدخلات بالأسلوب الذي يناسب طبيعة أساليب التنقيب في البيانات. وفي هذه المرحلة، يتم تهيئة بيانات التدريب والتعرف على الأعوام المحرفة وغير المحرفة للشركات محل الدراسة، ثم يتم تطبيق الخوارزمية للحصول على القواعد (Rules). وتتكرر هذه العملية حتى يتم توليد القواعد النهائية. وتطبق هذه القواعد على مجموعة بيانات فترة التدريب وفترة الاختبار وبالتالي يتم حساب النتائج بدقة. وتقوم خوارزمية سرب الجسيمات (PSO) على أساس استبدال أضعف النتائج في كل دورة حسب مقياس الجودة التي تقدمها خوارزمية (SVM) حتى يتم التوصل إلى القيمة المثلى (fitness function) وذلك بعدَ العديد من الدورات. وانتهت هذه الخطوة ببناء نموذج التنقيب في البيانات يستخدم التنبؤ بالغش في القوائم المالية في عينة فترة الاختبار.

3- اختبار أسلوب التنقيب في البيانات لعينة فترة الاختبار ومقارنة النتائج

بعد أن تم تدريب الخوارزمية من خلال بيانات فترة التدريب وبناء نموذج التنقيب في البيانات. يتم في هذه المرحلة اختبار أسلوب التنقيب في البيانات بمقارنة نتائج التصنيف وفق الخوارزميات التي تم استخدامها في التنقيب في البيانات مع نتائج التصنيف وفق النموذج الاحصائي التقليدي التي ظهرت نتائجه في الجدول رقم (8) و(9). ولكن قبل عرض نتائج أسلوب التنقيب باستخدام خوارزمية SVM-PSO، ولأغراض إضفاء جانب الدقة والموثوقية في عرض نتائج البحث الحالي، فقد قام الباحث باختبار النموذج أولاً على شركة عالمية أثارت ضجة عالمية إثرَ الفضيحة التي هزت بسمعتها نتيجة تضخيم أرباحها بمقدار (1.22) مليار دولار وعدم اكتشاف ذلك من خلال إجراءات التدقيق والنماذج الاحصائية التقليدية. وفيما يلي جدول رقم (10) يوضح نتائج تصنيف شركة توشيبا ل (6) أعوام قبل ظهور الغش في العام السابع من خلال نموذج BeneishMscore. والشكل رقم (3) يوضح نتائج التصنيف وفق نموذج التنقيب في البيانات الذي تم بناؤه باستخدام خوارزمية SVM-PSO:

الجدول رقم (10) تصنيف شركة توشيبا وفق نموذج beneish

ID\Index	2008	2009	2010	2011	2012	2013	2014
مؤشر DSRI	0.896	0.939	1.16	0.922	1.227	1.105	0.964
مؤشر GMI	1.015	1.288	0.856	0.975	0.978	0.984	0.973
مؤشر AQI	0.895	1.089	1.004	1.027	1.085	1.017	0.973
مؤشر SGI	1.079	0.883	0.963	1.023	0.956	0.955	1.135
مؤشر DEPI	0.769	0.903	1.119	1.082	0.995	1.344	1.326
مؤشر SGAI	1	1.057	0.905	0.939	1.049	1.015	0.984
مؤشر TATA	-0.042	0.003	-0.083	-0.07	-0.058	-0.022	-0.046
مؤشر LVGI	0.989	1.123	0.898	0.96	1.024	0.995	0.978
M-score	-2.76	-2.50	-2.77	-2.83	-2.58	-2.49	-2.59
classification	nofraud	nofraud	nofraud	nofraud	nofraud	nofraud	nofraud

ExampleSet (Apply Model)

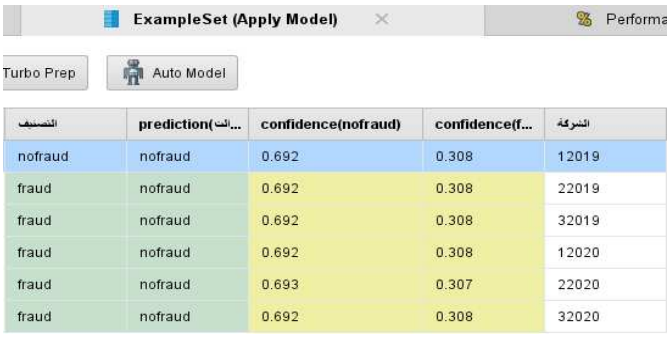
Auto Model

prediction(ات...	confidence(nofraud)	confidence(fraud)	العام
nofraud	0.625	0.375	2008
nofraud	0.625	0.375	2009
nofraud	0.624	0.376	2010
nofraud	0.598	0.402	2011
nofraud	0.548	0.452	2012
nofraud	0.534	0.466	2013
fraud	0.486	0.514	2014

الشكل رقم (3) نتائج تصنيف شركة توشيبا وفق نموذج التنقيب في البيانات

يُلاحظ أن نموذج Beneish M Score الذي أعطى نتيجة "nofruad" في جميع الأعوام بأنه لم يعطِ تنبؤ صحيح عن عملية التلاعب التي قامت بها شركة توشيبا عام (2014)، حيث ظهرت النتيجة (nofruad) في عام (2014) بينما في الواقع هي (fraud). أما نموذج التنقيب في البيانات–الشكل رقم (3)– فقد أعطى نتيجة مختلفة وهي (fraud)؛ أي أنه تنبأ بشكل صحيح على أنه يوجد "غش" تم فعلاً من قبل إدارة شركة توشيبا أدى إلى هز سمعة الشركة.

أما من ناحية الشركات الصناعية السورية محل الدراسة، فقد قام الباحث بتصنيف عينة فترة الاختبار عن عامي (2019) و(2020) باستخدام نموذج التنقيب في البيانات الذي تم استخدامه في مثال شركة توشيبا في برنامج التنقيب في البيانات باستخدام خوارزمية SVM-PSO كما هو مبين في الشكل رقم (4). وبمقارنة نتائج تصنيف شركة عينة الاختبار وفق النموذج الاحصائي – جدول رقم (8) و(9) – مع نتائج نفس الشركات لنفس الفترات وفقاً لنموذج التنقيب في البيانات، وجد الباحث عدم تطابق في (5) مشاهدات، هي الشركات رقم (22019,32019,12020,22020,32020). حيث أنه تم تصنيف هذه الشركات على أنها "محرفة" وفق نموذج beneish M score، في حين أنها "غير محرفة" وفقاً لنموذج التنقيب في البيانات.


ExampleSet (Apply Model) Performa

Turbo Prep Auto Model

التصنيف	prediction(الت...	confidence(nofraud)	confidence(f...	الشركة
nofraud	nofraud	0.692	0.308	12019
fraud	nofraud	0.692	0.308	22019
fraud	nofraud	0.692	0.308	32019
fraud	nofraud	0.692	0.308	12020
fraud	nofraud	0.693	0.307	22020
fraud	nofraud	0.692	0.308	32020

الشكل رقم (4) نتائج التصنيف لعينة الاختبار وفق نموذج التنقيب في البيانات

ومن خلال ما سبق؛ يتضح أنه يوجد اختلاف واضح بين نتائج نموذج beneish M score كأسلوب تقليدي في اكتشاف الغش ونتائج نموذج التنقيب في البيانات بنسبة 83.33% (5 اختلافات ÷ 6 مشاهدات) وهو اختلاف كبير، وبالتالي فإننا نرفض الفرضية الصفرية الأولى H01 التي تنص على أنه "لا يوجد اختلاف بين استخدام المدقق لأساليب التدقيق التقليدية وأساليب التنقيب في البيانات في اكتشاف الغش" ونقبل الفرضية البديلة H1؛ بأنه يوجد اختلاف بين استخدام المدقق لأساليب التدقيق التقليدية وأساليب التنقيب في البيانات في اكتشاف الغش.

ولا شك أنَّ النتائج التي أعطاها نموذج التنقيب في البيانات هي أكثر دقة، حيث أنه من خلال الواقع العملي واطلاع الباحث على تقارير التدقيق الصادرة عن مدققي الحسابات المعتمدين لتدقيق تلك الشركات، لم يُلاحظ وجود أي تحفظ على قوائم شركات الدراسة المدرجة في سوق دمشق للأوراق المالية[2]، وأن القوائم المالية فعلاً غير محرفة ولا يوجد فيها غش كما بيّن ذلك نموذج التنقيب في البيانات، على عكس نموذج Beniesh M Score الذي أظهر وجود غش بينما هو في الواقع لا يوجد غش. وهذا ما تم ملاحظته أيضاً في حالة شركة توشيبا، بأنَّ نموذج Beneish M score أعطى نتيجة "عدم وجود غش" بينما هو في الواقع "يوجد غش"، أي أنَّ النموذج وقع في خطر " القبول الخاطئ" وهو خطر فاعلية؛ أي أنه قَبِل البيانات التي أصدرتها شركة توشيبا، في حين تبين لاحقاً على أنَّ هذا القبول خاطئ إثر الفضيحة التي تورطت بها شركة توشيبا والتي هي واحدة من أكبر الشركات اليابانية والتي لها صدىً عالمي (Edwards, 2015). وهذا ما يدعم فاعلية التدقيق عند استخدام نموذج التنقيب في البيانات، وبالتالي فإننا نرفض الفرضية الصفرية الثانية H02 التي تنص على أنه" لا يؤدي استخدام المدقق لأساليب التنقيب في البيانات إلى زيادة فاعلية اكتشاف الغش في القوائم المالية" ونقبل الفرضية البديلة H2 بأن استخدام المدقق لأساليب التنقيب في البيانات أدى إلى زيادة فاعلية اكتشاف الغش في القوائم المالية مقارنةً باستخدام نموذج Beneish M Score.

النتائج والتوصيات

1) النتائج:

1- يوجد اختلاف في النتائج التي تم التوصل إليها عند استخدام المدقق لأساليب التدقيق التقليدية في اكتشاف الغش وبين استخدامه لأساليب التنقيب في البيانات. حيث لم تتطابق نتائج النموذجين التقليدي والتنقيب في البيانات بنسبة اختلاف قدرها (83.33%) في ضوء مجتمع الدراسة.

2- تفوّق أساليب التنقيب في البيانات (خوارزمية سرب الجسيمات PSO مع خوارزمية آلات متجهات الدعم SVM) في التنبؤ بغش الإدارة على النماذج الاحصائية التقليدية التي تُستخدم للتنبؤ بحالات الغش.

[2] http://www.dse.gov.sy/

3- فشل نموذج Beneish M score في التنبؤ في حالة الغش التي حصلت عام (2014) لإحدى الشركات العالمية التي تعمل في المجال الصناعي التقني(شركة توشيبا). ونجاح نموذج التنقيب في البيانات في اكتشاف غش وتلاعب إدارة شركة توشيبا.

2) التوصيات:

1- إجراء المزيد من الدراسات على مجتمع إحصائي أكبر من الذي أجرته هذه الدراسة، والاسترشاد بنتائج هذه الدراسة للقيام بمزيد من الدراسات المستقبلية حول دور أساليب التنقيب في البيانات في تعزيز فاعلية التدقيق بالتطبيق على مجتمع إحصائي أكبر من ثلاث شركات وعدد أكثر من الفترات المالية.

2- ضرورة تنمية القدرات المهنية، والمهارات الفنية للمدققين من خلال القيام بدورات تدريبية في استخدام أساليب التنقيب في البيانات تحت اشراف متخصصين في مجال نظم المعلومات والبرمجة.

3- تعاون كلية الاقتصاد مع كلية الهندسة المعلوماتية في تصميم برامج تستخدم خوارزميات معينة من شأنها أن تساعد المدققين في تحليل البيانات واستخراج المعارف والأنماط منها، لا سيما في ضوء التطورات المتسارعة في بيئة الأعمال واعتمادها على نظم تقنية وبيانات متشعبة.

المراجع References

أولاً: المراجع العربية:

1. الخياط، إياد.(8-3-2018).ما هو علم البيانات؟. استرجعت في تاريخ 15-3-2021 من https://datasciencearabi.com
2. الجبلي، وليد سمير عبد العظيم. (2020). أثر استخدام الانحدار اللوجستي أحد أساليب التنقيب في البيانات (Data Mining) في دعم الرأي المهني لمراجعي الحسابات – دراسة تطبيقية. مجلة الإسكندرية للبحوث المحاسبية، 4م (ع2)، ص ص 172-132
3. القاضي، حسين، دحدوح، حسين، وقريط، عصام،(2013) أصول المراجعة الجزء الأول. دمشق: سوريا. منشورات جامعة دمشق كلية الاقتصاد.ص:611.
4. أحمد محمد أبو الخير، أسامه. (2019). دور استخدام أساليب التنقيب في البيانات لتحسين تقديرات مراقب الحسابات في مدى وجود أخطاء جوهرية بالقوائم المالية: دراسة ميدانية في بيئة الأعمال المصرية. مجلة الدراسات التجارية المعاصرة، م5(ع7) ص ص 305-347
5. رزوق، راكان. (2013). التنقيب في البيانات الأسس النظرية والتطبيق. دمشق: المركز العربي للتعريب والترجمة والنشر. ص: 211.
6. مثنى، صبحي سليمان، رياض، محمود صالح. (2012). توظيف التحليل العنقودي وطريقة الجار الأقرب في التعرّف على الأنماط مع تطبيق على نوعية المياه الجوفية في محافظة نينوى. المجلة العراقية للعلوم الاحصائية، م12(ع21)، ص ص 165-141.

ثانياً: المراجع الأجنبية:

1. AICPA (2002),Consideration of Fraud in a Financial Statement Audit. Statement on Auditing SAS No.99
2. Appelbaum, D., Kogan, A., & Vasarhelyi, M. A. (2017). Big Data and analytics in the modern audit engagement: Research needs. Auditing: A Journal of Practice & Theory, 36(4), 1-27.
3. Breiman, L. (2001). Random forests. Machine learning, 45(1), 5-32.
4. Dawalekar, P., Gharat, N., Gaikwad, S., & Kundale, J. (2019). Data Mining Techniques for Smart Fitness. In 2nd International Conference on Advances in Science & Technology (ICAST).
5. Del Valle, Y., Venayagamoorthy, G. K., Mohagheghi, S., Hernandez, J. C., & Harley, R. G. (2008). Particle swarm optimization: basic concepts, variants and applications in power systems. IEEE Transactions on evolutionary computation, 12(2), 171-195.
6. Edwards Jim.(21-7-2015). Toshiba CEO resigns over massive accounting scandal Retrieved on 28-3-2021 from https://www.businessinsider.com/r-toshiba-ceo-resigns-over-massive-accounting-scandal-2015-7
7. Hoffman, V. B., & Zimbelman, M. F. (2009). Do strategic reasoning and brainstorming help auditors change their standard audit procedures in response to fraud risk?. The Accounting Review, 84(3), 811-837.
8. Hooda, N., Bawa, S., & Rana, P. S. (2018). Fraudulent firm classification: a case study of an external audit. Applied Artificial Intelligence, 32(1), 48-64.
9. http://www.dse.gov.sy /
10. http://www.dse.gov.sy/user/?page=law_show&id=25&pr=R

11. International Federation of Accountants (IFAC) (2020). " The Auditor's Responsibilities. Relating to Fraud in an Audit of Financial Statements", International Standards on Auditing ISA 240.
12. Jakkula, V. (2006). Tutorial on support vector machine (svm). School of EECS, Washington State University, V37.pp1-13
13. MacCarthy, J. (2017). Using Altman Z-score and Beneish M-score models to detect financial fraud and corporate failure: A case study of Enron Corporation. International Journal of Finance and Accounting, 6(6), 159-166.
14. McCluskey, W. and Anand, S. (1999)."The application of intelligent hybrid techniques for the mass appraisal of residential properties", Journal of Property Investment & Finance, Vol. 17 No. 3, pp. 218-239 .
15. Mollah, M. D., & Sakib, I. A. (2020). Detection of Financial Statement Frauds Using Beneish Model: Empirical Evidence from Listed Pharmaceutical Companies in Bangladesh. International Journal of Management, Accounting and Economics, 7(9), 506-521.
16. Nageswari, S., Goel, Pallavi..(2019). Comparison of Classification Techniques on Data Mining. international Journal of Emerging Technology and Innovative Engineering, Volume 5, Issue 5.pp1-6.p1
17. Nawaiseh, A. K., Abbod, M. F., & Itagaki, T. (2020). Financial Statement Audit using Support Vector Machines, Artificial Neural Networks and K-Nearest Neighbor: An Empirical Study of UK and Ireland. International Journal of Simulation Systems, Science & Technology, 21(2).71-76
18. Poli, R., Kennedy, J., & Blackwell, T. (2007). Particle swarm optimization. Swarm intelligence, 1(1), 33-57.
19. Repousis, S. (2016). Using Beneish model to detect corporate financial statement fraud in Greece. Journal of Financial Crime,(23)(4),1063-1073.
20. Sharma, A., & Panigrahi, P. K. (2012). A review of financial accounting fraud detection based on data mining techniques. international Journal of computer applications,(39)(1),37-47.
21. Silwattananusarn, T., & Tuamsuk, K. (2012). Data mining and its applications for knowledge management: a literature review from 2007 to 2012. International Journal of Data Mining & Knowledge Management Process (IJDKP) Vol.2, No.5, 2012, pp. 13-24
22. Yan, J., Wang, X., Wang, B., & Zhang, Y. (2019). Research on application of data mining technology in risk assessment process of audit. In 2019 International Conference on Economic Management and Model Engineering (ICEMME) (pp. 487-491). IEEE.