مجلة البحوث المالية والتجارية
المجلد (٢٣) – العدد الرابع – أكتوبر ٢٠٢٢

تقييم دقة أساليب التنقيب في البيانات Data Mining (DM) في الحد من المخاطر الائتمانية وانعكاساها على جودة القوائم المالية بالقطاع المصرفي–دراسة حالة

Evaluating the accuracy of data mining (DM) methods in reducing credit risks and their impact on the quality of financial statements in the banking sector - a case study

د/ سوزي فاروق النقودي

مدرس محاسبة

كلية تكنولوجيا الادارة ونظم المعلومات - جامعة بورسعيد

E-mail: suzy_elnokoudy@yahoo.com

رابط المجلة: https://jsst.journals.ekb.eg/

المستخلص:

أجبرت الهزات الاقتصادية العالمية وعمليات الاندماج الكبرى في القطاع المصرفي وبصفة خاصة البنوك، على إعادة التفكير في تبني استراتيجيّات جديدة تتواكب مع تطور التكنولوجيا وظهور تقنيات الذكاء الاصطناعيAI، مما أدى إلى ضرورة إحداث تغيرات جذرية في القطاع المصرفي المصري والسعي نحو التحول إلى بنوك الجيل الجديد (البنوك الالكترونية) التي تتمتع بتكنولوجيا مصرفية جديدة وأساليب مبتكرة. وهو ما دفع البنك المركزي المصري إلى إنشاء إدارة تسمى "إدارة تحصيل مخاطر الائتمان" للإعلان عن التقييم الائتماني والتقرير الائتماني للمقترض لإثبات الجدارة الائتمانية للعميل، بالاعتماد على بعض الأساليب التكنولوجية والإحصائية الحديثة لتقليل المخاطر الائتمانية وبصفة خاصة مخاطر منح القروض، وتحسين جودة القوائم المالية والأداء المالي للبنك، وعلاقته بالعملاء، مما يقلل الوقت والتكلفة.

تهدف الدراسة إلى تقييم دقة الأساليب المختلفة للتنقيب في البياناتDM – باعتباره أحد التقنيات الحديثة في مجال الذكاء الاصطناعيAI– في التنبؤ بمنح القروض للحد من المخاطر الائتمانية وانعكاس ذلك على جودة القوائم المالية.

توصلت الدراسة إلى أن أساليب التنقيب في البيانات DM المختلفة تعمل على تحسين دقة التنبؤ وسرعة منح القروض المصرفية، وفقا لتصنيفاتها القائمة على التصنيف والتجميع والتنبؤ والتصور.

كما توصلت الدراسة إلى تباين مستوى دقة التنبؤ بين أساليب التنقيب في البيانات DM ، حيث أن نموذج الغابات العشوائية RF هو الأعلى دقة في التنبؤ وفقا لمنحنى ROC والأكثر حساسية يليه أسلوب شجرة القرار DT ويأتي في المرتبة الثالثة أسلوب الانحدار اللوجستي LR .

الكلمات الافتتاحية: التنقيب في البياناتDM، المخاطر الائتمانية، منح القروض، القوائم المالية.

Abstract:

Global economic crises and major mergers in the banking sector have forced the adoption of new strategies for developing technology and AI methods, which has led to the need for the Egyptian banking sector to transform into a new generation of banks (e-banks) that adopt new banking technology.

The Central Bank of Egypt has established a department called "the Credit Risk Collection Department" to calculate the credit score & credit report of the borrower to prove the creditworthiness of the customer, relying on some modern technological and statistical methods to reduce credit risks, in particular, the risks of granting loans, and improving the quality of the financial statements and the bank's financial performance, and its relationship with customers, which reduces time and cost.

This study aims to evaluate the accuracy of the different DM methods - as one of the modern AI techniques - in predicting the granting of loans to mitigate credit risks and its impact on the quality of financial statements.

The study found that different DM methods improve prediction accuracy and speed of granting bank loans, according to their classifications based on classification, aggregation, prediction, and visualization.

The study also found a variation in the level of prediction accuracy between the DM methods, the higher prediction accuracy is the RF model according to the ROC curve and the most sensitive, and then the DT method and the least accuracy is the LR method.

Keywords: Data mining DM, credit risk, granting loans, financial statements.

١/ الإطار العام للدراسة

١/١ مقدمة البحث:

في عالم اليوم، يلعب القطاع المصرفي دورًا رئيسيًا في التنمية المالية لأي بلد، بشكل عام، يحتوي القطاع المصرفي على العديد من أنواع البيانات التاريخية في مجموعة من قواعد البيانات الضخمة غير المتجانسة، وفي ظل توجه الدولة نحو التحول الرقمي، وتوجه البنك المركزي المصري نحو تبني جيل جديد من البنوك وهو البنوك الالكترونية ، فإن الأمر يتطلب ضرورة الوصول إلى طريقة أفضل لاستخدام تلك البيانات تواكباً مع التكنولوجية الحديثة التي تغزو القطاع المصرفي على مستوى العالم.

لذلك ، تطلبت المنافسة المتزايدة لتغيرات السوق استخبارات مصرفية لتحليل تلك البيانات الهائلة، من خلال زيادة التوجه نحو تقنيات وتطبيقات الذكاء الاصطناعي، حيث نجحت بنوك الجيل الجديد (البنوك الالكترونية) ذات التكنولوجيا المصرفية العالية والأساليب المبتكرة لإدارة أعمالها، في التوجه نحو إنشاء قاعدة بيانات ضخمة وشاملة لاستخدامها في مجال التخطيط الاستراتيجي من أجل البقاء وتحقيق الميزة التنافسية، والوفاء بتوقعات العملاء، مما أدى إلى تعريضها بصورة أكبر إلى العديد من المخاطر، وظهور نوع جديد من المخاطر وهي المخاطر الالكترونية، وزيادة التوجه نحو إيجاد أساليب تكنولوجية حديثة للحد من مخاطر الائتمان التي تؤثر بشكل جوهري على المركز المالي للبنك ، حيث أصبح للعميل فرصة أعظم لاختيار البنوك حول العالم ذات المراكز المالية العالية والمستقرة لاستثماراته.

في أواخر الثمانيات ظهر التنقيب في البيانات DM كأحد الحلول الناجحة لتحليل كميات ضخمة من البيانات، وذلك بتحويلها من مجرد معلومات متراكمة وغير مفهومة (بيانات) إلى معلومات قيمة يمكن استغلالها والاستفادة منها، والتي تقوم على (التصنيف، التجميع، التنبؤ، والتصوير) من خلال استخدام بعض الأساليب كالشبكات العصبية المتطورة، ونموذج الانحدار اللوجستي، وشجرة القرار ، والغابات العشوائية، والخوارزميات الجينية. الخ.

وقد ذكر المعهد الأمريكي للمحاسبين القانونيين Association of International Certified Professional Accountants (AICPA) بأن التنقيب في البيانات DM هي واحدة من أعلى عشرة تكنولوجيات مستخدمة مستقبلاً، بينما وضعها معهد المراجعين الداخليين من

الأولويات الأربعة للبحث (kirkos & Manopoulos , 2007) ، وذلك لقدرتها على تحليل البيانات، وإنشاء نماذج وصفية وتنبؤية فى آن واحد، وهو ما تراه الباحثة ضرورة للوفاء بمجهودات القطاع المصرفي ، والذي يسعى جاهداً إلى خلق وابتكار وتقديم خدمات مصرفية جديدة، محاولاً بذلك كسب رضا العملاء واستمالتهم بغرض نيل ثقتهم وبالتالي زيادة معاملاته معهم.

٢/١ مشكلة البحث:

تعد مواجهة المخاطر الائتمانية للبنوك من الأولويات التي تحظى بأولوية بحثية كبيرة في الفترة الحالية، في ظل المنافسة الحادة بين البنوك، للحفاظ على عملائها واجتذاب عدد كبير من العملاء، من أجل تحسين كفاءة الأنشطة المصرفية، وتحسين القدرة التنافسية، والحد من المخاطر المصرفية، وزيادة جودة القوائم المالية.

يعتبر منح القروض واحدة من أهم وأصعب المخاطر الائتمانية التي يتعرض لها القطاع المصرفي، نظراً لحساسية القرار وتأثيره على الوضع المالي في حالة عدم القدرة على السداد في الموعد المحدد. وفي ظل الأزمات المتتالية التي يواجهها العالم، مثل أزمة Covid-19 واضطرار البعض إلى ترك أعمالهم والجلوس في المنزل، فقد زاد الطلب على الاقتراض من القطاع المصرفي لمواجهة الأزمات المالية والغلاء الذي يسود العالم، وزاد معه بالتالي الخطورة المالية التي يتعرض لها القطاع المصرفي، نتيجة عدم سداد القرض أو التأخر في السداد، أدرك القطاع المصرفي أن عليه وبشكل سريع اتخاذ القرارات الاستراتيجية الصحيحة بسرعة..

ومع التطورات التكنولوجية المتلاحقة في مجال نظم المعلومات في القطاع المصرفي، وتوجه الدولة نحو التحول الرقمي، يسعى البنك المركزي المصري الى تحويل البنوك المصرية الى بنوك الكترونية، وتواكباً مع ذلك قام البنك المركزي المصري بتحديث نظام التسجيل الائتماني للعملاء، حتى يكون أكثر فاعلية وشفافية، وبالتالي رفع كفاءة الجهاز المصرفي، ومقدرته على إدارة المخاطر، بما يساهم على تقوية المركز المالي للبنوك المصرية والأجنبية العاملة في مصر لمواكبة التغيرات التي تحدث في هيكل الأنشطة الاقتصادية على المستوى الدولي.

قام البنك المركزي المصري باستحداث " *إدارة تجميع مخاطر الائتمان*"، بهدف توفير قاعدة بيانات مجمعة تتضمن كافة البيانات والمعلومات الايجابية والسلبية عن العملاء، وأطرافهم المرتبطة المنتظمين وغير المنتظمين في السداد، بغرض دعم البنوك المانحة للقروض

بالاطلاع على تلك البيانات المجمعة قبل إقرار منح الائتمان أو زيادته أو تجديده عبر شبكة معلومات البنك المركزي (البنك المركزي ، ٢٠٢١).

تقوم تلك الادارة بإعداد كلا من التقرير الرقمي Credit Report، والتقييم الرقمي Credit Score للعملاء، والتقرير الرقمي (هو مستند رسمي موجود فيه كل بيانات التسهيلات الائتمانية الخاصة بكل فرد مثل: عدد القروض ومدى الالتزام بسداد الأقساط، عدد أيام التأخير والمبالغ المتأخرة وغيره)، والتقييم الرقمي (هو مؤشر يعكس الموقف الائتماني للعميل ويتم احتسابه بطرق احصائية متطورة أخذاً فى الاعتبار عوامل مثل الدخل ، السن ، الحالة الاجتماعية ، مدى التزامه فى السداد ... الخ)، ويعكس التقييم الرقمي Credit Score درجة المخاطرة للعميل، وكلما زادت درجة التقييم الرقمي للعميل كلما اعتبر العميل ذو درجة مخاطر منخفضة والعكس صحيح (البنك المركزي ، ٢٠٢١).

وفي هذا الشأن وجدت الباحثة حاجة ملحة إلى دراسة وتقييم استخدام أساليب "التنقيب في البيانات DM" المختلفة في القطاع المصرفي، في التنبؤ بقدرة العميل على سداد القرض في الموعد المحدد للتقليل من مخاطر الائتمان الناتجة عن القروض المتعثرة، باعتبار التنقيب في البيانات DM أحد تطبيقات الذكاء الاصطناعي التي تساعد على التنبؤ ودقة اتخاذ القرار، والقائمة على تحليل البيانات وإنشاء نماذج وصفية وتنبؤية فى آن واحد، وذلك لرفع جودة القوائم المالية التي تتأثر في الكثير من بنودها بمؤشرات المخاطر الائتمانية.

ومما سبق يمكن إيجاز مشكلة البحث في التساؤلات البحثية التالية:

أولاً: ما المقصود بالتنقيب في البيانات DM ؟
ثـانيـاً: ما هي المخاطر التي يواجهها القطاع المصرفي ؟
ثـالـثـاً: ما هو تأثير المخاطر الائتمانية المصرفية على جودة القوائم المالية ؟
رابـــعـاً: ما مزايا استخدام أساليب التنقيب في البيانات للحد من المخاطر الائتمانية ؟
خامساً: هل يوجد اختلاف في درجة دقة التنبؤ بقرارات منح القروض بين الأساليب المختلفة للتنقيب في البيانات DM ؟

٣/١ أهمية البحث:

تنبع أهمية البحث من تحول الدولة بصفة عامة إلى التحول الرقمي، واستخدام التكنولوجيا في كافة قطاعاتها، وسعى البنك المركزي المصري نحو ممارسة الخدمات البنكية عبر المواقع

الالكترونية للبنك، والتحول نحو البنوك الالكترونية. فقد أصبح استخدام تطبيقات الذكاء الاصطناعي AI ضرورة تسعى البنوك إلى مواكبتها، حيث يساعد على تحليل وتصنيف واستكشاف البيانات والدلالات وأنماط سلوكيات العملاء من سجلات القطاع المصرفي، للتنبؤ بأنماط وتوجهات العملاء في المستقبل، وبالتالي فإن التوجه نحو الاستفادة من أساليب التنقيب في البيانات DM باعتبارها أحد تطبيقات الذكاء الاصطناعي AI، ربما يساعد على الحد مخاطر الائتمان (بصفة خاصة التنبؤ بمنح القروض) بحيث يمكن معالجتها في أسرع وقت ممكن ، وتجنب المخاطر الائتمانية والتي تؤثر على جودة القوائم المالية بشكل سلبي وبالتالي قدرته على تحقيق ميزة تنافسية.

وتنقسم إلى:

١/٣/١ الأهمية العلمية:

تنبع أهمية البحث العلمية من تأكيد المؤسسات المهنية الرئيسية في مجال المحاسبة على أهمية التنقيب في البيانات DM، فقد أشار المعهد الأمريكي للمحاسبين القانونين المعتمدين American Institute of Certified Public Accountants (AICPA) أن التنقيب في البيانات DM هو واحد من أفضل عشر تكنولوجيات في المستقبل، كما وضعها معهد المراجعين الداخليين من الأولويات الأربعة للبحث (kirkos & Manopoulos , 2007) ، ونظراً لقلة البحوث والدراسات المتخصصة – في حدود علم الباحثة – التي تناولت أثر استخدام التنقيب في البيانات DM في القطاع المصرفي– من وجهة النظر المحاسبية– فإن الباحثة تسعي إلى تقييم أثر استخدام الأساليب المختلفة للتنقيب في البيانات على تحليل بيانات العملاء المحفوظة في البنوك، للتنبؤ بتوجهات العملاء وربحيتهم، وقدرتهم على سداد القروض، مما قد يساهم في تخفيف مخاطر الائتمان المصرفي، وبالتالي ينعكس إيجابيًا جودة القوائم المالية.

٢/٣/١ الأهمية العملية :

تنبع أهمية البحث العملية من خلال تحديد السرعة والدقة والمنفعة التي تعود على العملاء، والقطاع المصرفي ممثلاً في البنوك، من التوجه نحو استخدام تقنيات التنقيب في البيانات DM بدلاً من الطرق والأساليب الاحصائية التقليدية، للحد من المخاطر الائتمانية الناتجة عن منح القروض، والتنبؤ بتعثر العملاء عن السداد، من خلال تصنيف العملاء وأرباحهم، وسجلات الائتمان، ومساعدة إدارات الائتمان على اتخاذ القرارات الصحيحة والفعالة

عند منح القرض لعملائها، وتمكين الإدارة في تحديد السياسات والإجراءات ذات الكفاءة العالية نحو الأمان والاستقرار المصرفيين.

٤/١ هدف البحث:

يهدف هذا البحث إلى إعطاء منظور جديد حول كيفية جعل عملية منح القروض أكثر دقة وكفاءة وسرعة، من خلال تقييم أكثر أساليب التنقيب في البيانات DM دقة في التنبؤ (بالتركيز على أسلوب الانحدار اللوجستي LR ، وشجرة القرار Decision Tree ، والغابات العشوائية RF) باعتبارها أكثر الأساليب التي دعمتها الدراسات المختلفة في قدرتها على التنبؤ بمنح القروض، وبالتالي الحد من المخاطر الائتمانية للبنوك، مما يزيد من الوضع التنافسي الجيد للبنوك نتيجة زيادة جودة القوائم المالية التي يعتمد عليها المستثمرين.

٥/١ منهجية البحث:

استخدام المنهج التحليلي الوصفي باستخدام تقنيات التنقيب DM عن البيانات واستخدام خوارزمية التصنيف والتجميع في تحليل البيانات واكتشاف المعرفة والتنبؤ بالبيانات عن طريق الأساليب المختلفة للتنقيب في البيانات DM مثل أسلوب الانحدار اللوجستي LR ، وشجرة القرارDT، والغابات العشوائيةRF. وقد قامت الباحثة بإجراء دراسة حالة على بيانات أحد البنوك التي تم الحصول عليها من موقع www.kaggle.com

٦/١ خطة البحث:

لتحقيق هدف البحث، تم تقسيم البحث (بخلاف المقدمة) إلى المحاور التالية:

١/ الاطار العام للدراسة

٢/ الاطار النظري : ويشتمل على :

١/٢ الدراسات السابقة.

٢/٢ التنقيب في البيانات Data Mining (DM) كأحد أدوات الذكاء الاصطناعي.

٣/٢ المخاطر التي يواجهها القطاع المصرفي.

٤/٢ المخاطر الائتمانية وانعكاساتها على جودة القوائم المالية للبنوك.

٢/٥ دور التنقيب في البياناتDM في الحد من المخاطر الائتمانية في القطاع المصرفي مما يؤثر على جودة القوائم المالية.

٢/٦ أساليب التنقيب في البياناتDM المستخدمة في التنبؤ بمنح القروض في القطاع المصرفي.

٣/ الدراسة التطبيقية

٤/ الخلاصة والنتائج والتوصيات والأبحاث المستقبلية

٥/ المراجع

٢/ الإطار النظري

٢/١ الدراسات السابقة:

٢/١/١ الدراسات المتعلقة بالتنقيب في البيانات للحد من مخاطر الائتمان:

- دراسة (GU, 2022) بعنوان:

"Data-Driven Fuzzy Multiple Criteria Decision Making and its Potential APplications-2021"

هدفت هذه الدراسة إلى استخدام تحليل نموذج آلات ناقلات الدعم (SVM)، ونموذج الانحدار الخطي، والنموذج اللوجستي، لتجنب المخاطر المالية المنتظمة من التعاملات المالية على الإنترنت، وتحسين كفاءة قدرة شركات الخدمات القانونية على تقييم مخاطر الائتمان .

قدمت هذه الدراسة نموذجًا جديدًا في مجال تقييم مخاطر الائتمان للشركات، وهو نموذج قائم على تعدد النماذج لتقييم وتحليل الائتمان، ومستوى مخاطر الشركات المدرجة، وصنع القرار المتكامل. في النموذج المتكامل المقترح، يصنف نموذج SVM عينة البيانات من الشركات المدرجة، ويستخدم النموذج اللوجستي لتحليل الانحدار في تقييم مخاطر الائتمان. يتم تطبيق نموذج FMCDM المقترح على أساس مجموعة ضبابية Fuzzy للحصول على مؤشرات التقييم، ثم استخدام عملية التسلسل الهرمي التحليلي (AHP) للحصول على وزن نسبي للفهارس. أخيرًا ، قامت الدراسة بإجراء تحليل الملاءمة وفقًا لحالة المخاطر الحالية لشركة العينة ونتائج حالة المخاطر للنموذج المتكامل المقترح.

وتوصلت الدراسة أن تقنية استخراج البيانات القائمة على نموذج SVM-Logistic + FMCDM لديها دقة عالية وتطبيق عملي في تقييم مخاطر الائتمان من شركات الخدمات القانونية.

-دراسة (Pandimurugan, et al., 2022) بعنوان:

Random forest tree classification algorithm for predicating " "loan

هدفت هذه الدراسة إلى تحليل استخدام نموذج التعلم الآلي Machine Learning اعتماداً على خوارزمية الغابة العشوائية للتنبؤ بالموافقة على منح القرض لمقدم الطلب من عدمه، وبالتالي تقليل المخاطر التي ينطوي عليها الأمر بشكل فعال في المؤسسات المالية أو البنوك.

وتوصلت الدراسة أن خوارزمية الغابات العشوائية لها أداء ممتاز في التنبؤ الافتراضي بمنح القروض أفضل من الانحدار اللوجستي، وذلك بهدف استهداف واكتساب عملاء جدد بشكل أفضل .

- دراسة (Ernawati ,2021) بعنوان :

A review of data mining methods in RFM-based customer " "segmentation

هدفت هذه الدراسة إلى تحليل طرق التنقيب في البيانات DM التي تتعاون مع نموذج RFM وتوليها اقتراح إطار عمل لتجزئة العملاء. قامت الدراسة بمراجعة شاملة للأدبيات المنشورة في ٢٠١٥-٢٠٢٠.

توصلت الدراسة إلى أن الطرق الأكثر استخدامًا هي التجميع والتصور من الطرق السبع لـ DM التي تم تحليلها، نظرًا لزيادة وظيفة التصور والحاجة إلى أخذ البيانات الجغرافية والديموغرافية للعملاء في الاعتبار في التحليل ، قدمت الدراسة إطارًا جديدًا لاستخدام طرق DM مع التجزئة القائمة على RFM في بيئة أنظمة المعلومات الجغرافية (GIS)، يساعد إطار العمل هذا المحللين على استخدام طرق DM للكشف عن خصائص العملاء وفهمها، بحيث يمكن للشركات تحديد السوق المستهدف وتطوير استراتيجية تسويق لزيادة ميزتها التنافسية.

- دراسة (Al-Hashedi , 2021) بعنوان :

"Financial fraud detection aPplying data mining techniques: A comprehensive review from 2009 to 2019"

هدفت هذه الدراسة إلى تقديم مراجعة شاملة لأحدث الأبحاث في الكشف عن الاحتيال المالي من عام ٢٠٠٩ إلى عام ٢٠١٩، وتصنيفها بناءًا على أنواع الاحتيال، وتقنية التنقيب في البيانات المستخدمة في اكتشاف الاحتيال المالي.

توصلت الدراسة من خلال تصنيف المقالات إلى أربع مجموعات رئيسية (الاحتيال المصرفي، والاحتيال في التأمين، والاحتيال في البيانات المالية، والاحتيال بالعملات المشفرة)، أنه تم استخدام ٣٤ تقنية للتنقيب عن البيانات لتحديد الاحتيال عبر التطبيقات المالية المختلفة. تم ملاحظة أن أسلوب آلات ناقلات الدعم SVM هي واحدة من أكثر تقنيات الكشف عن الاحتيال المالي استخدامًا والتي تحمل حوالي ٢٣ ٪ من الدراسة الإجمالية، تليها كل من أسلوب بايز NaIve Bayes , شجرة القرارات Random Forest

٢/١/٢ الدراسات المتعلقة بانعكاس مخاطر الائتمان على القوائم المالية:

- دراسة (Peng, & et.al., 2022) بعنوان:

"Bank Financial Risk Prediction Model Based on Big Data"

هدفت هذه الدراسة إلى استخدام التنبؤ بالمخاطر المالية كأسلوب منهجي للتنبؤ بالمخاطر غير المتوقعة في الأنظمة المصرفية. حيث تتطلب القضايا التي تنطوي على سوء التوقيت والدقة المنخفضة في طرق التنبؤ بالمخاطر الحالية طريقة فعالة للتنبؤ بالمخاطر. كما هو الحال عند استخدام البيانات الضخمة في مختلف المجالات، تلعب التكنولوجيا دورًا مهمًا في الخدمات المالية ويمكن استخدامها للتنبؤ بدقة وفي الوقت المناسب بإمكانيات المخاطر.

اقترحت الدراسة طريقة هجينة فعالة للتنبؤ بشكل مناسب وفعال بالمخاطر المالية في الأنظمة المصرفية. باستخدام خوارزميات Lasso والانحدار الخطي معاً عبر مميزات البيانات الكبيرة وتقنيات الإطار. وبإضفاء الطابع الرسمي المناسب على مشاكل المخاطر المالية للبنك، يتم الحصول على بيانات المخاطر ومعالجتها.

-دراسة (Hassoon, et al.,2022) بعنوان :

"The effect of credit risk on financial performance in Iraqi private commercial banks: Bank of Baghdad as a model"

هدفت هذه الدراسة إلى تحليل مخاطر الائتمان في أحد البنوك العراقية الخاصة، ومعرفة كيفية تأثير هذه المخاطر على الأداء المصرفي لبنك.

توصلت الدراسة إلى أن الاهتمام بعملية تحليل المعلومات والبيانات الخاصة بحالة العميل سيعطي لإدارة مخاطر الائتمان تصورًا جيدًا حول مخاطر الائتمان التي تصاحب هذه عملية الائتمان، كما أن عمليات التأمين ضد مخاطر الائتمان من وسائل التنمية الجديدة التي تنشأ بين شركة التأمين والبنوك كنتيجة جيدة يفرضها التطور التكنولوجي والعولمة المالية.

- دراسة (الفارسي وآخرون، ٢٠٢٢) بعنوان :

"أسلوب التنقيب في البيانات كأحد ابتكارات تكنولوجيا المعلومات في تحسين جودة التقارير المالية -دراسة ميدانية"

هدفت الدراسة إلى تقديم إطار مقترح لأثر استخدام نموذج التنقيب في البيانات علي الكشف عن حالات الفساد المالي والاداري بوزارة الإعلام الكويتية.

وتوصلت الدراسة إلي النتائج التالية: يستخدم في التنقيب في البيانات عدة طرق تصنيف مستمدة من مجالات الإحصاء والذكاء الاصطناعي، توجد ثلاث طرق تتمتع بسمعة طيبة لقدرات التصنيف، هذه الأساليب هي أشجار القرار، والشبكات العصبية وشبكات العقد بايزي. يتم الكشف عن عمليات الفساد المالي والإداري باستخدام أساليب التنقيب في البيانات التي تجعل المراجعة أسهل من خلال تنظيم وتحليل البيانات بطريقة أكثر كفاءة وفعالية. تم اختبار الفرض من خلال اختبار "ت" لعينة واحدة one sample t-test وتم قبول الفرض البديل بوجود اتفاق تو دلالة معنوية بين أراء المستقصي منهم حول محور (مزايا تفعيل آليات أسلوب التنقيب في البيانات بوزارة الإعلام الكويتية). باختبار الفرض من خلال اختبار "T" لعينة واحدة. وتم قبول الفرض البديل بوجود اتفاق ذو دلالة معنوية بين آراء المستقصي منهم حول محور مهام ومراحل التنقيب في البيانات بوزارة الإعلام الكويتية.

- دراسة (حسين، ٢٠٢١) بعنوان:

" أثر الافصاح عن مخاطر الائتمان المصرفي على تحسين الأداء المالي للبنوك التجارية في بيئة الأعمال المصرية"

هدف الدراسة إلى التعرف على تأثير الإفصاح عن المخاطر الائتمانية على تحسين الأداء المالي للبنوك التجارية في بيئة الأعمال المصرية، وذلك من خلال قياس الإفصاح عن المخاطر الائتمانية من خلال المؤشرات التالية: نسبة القروض المتعثرة، نسبة مخصصات خسائر القروض، نسبة القروض والسلفيات للموارد الخارجية والذاتية، نسبة حقوق المساهمين إلى إجمالي القروض، نسبة السيولة، نسبة الرافعة المالية، نسبة الكفاءة المالية، نسبة كفاية رأس المال، وحجم البنك، وتم قياس الأداء المالي من خلال المؤشرات التالية: معدل العائد على الأصول ROA، ومعدل العائد على حقوق الملكيةROE ، معدل العائد على السهمEPS ، ولتحقيق أهداف البحث تم الاعتماد على المنهج الوصفي والتحليلي باستخدام العديد من الأساليب الإحصائية مثل الإحصاء الوصفي

توصلت الدراسة إلى وجود أثر معنوي ذو دلالة إحصائية بين الإفصاح عن المخاطر الائتمانية المصرفية وتحسين الأداء المالي للبنوك التجارية في بيئة الأعمال المصرية.

- دراسة (عبد الغفار، ٢٠٢٠) بعنوان:

"استخدام أسلوب التنقيب في البيانات DM لدعم المحتوى المعلوماتي للقوائم المالية المستقبلية وأثر ذلك على تعزيز كفاءة القرارات الاستثمارية في السوق المالي المصري – دراسة تطبيقية "

هدفت هذه الدراسة إلى دراسة وتحليل استخدام أسلوب التنقيب في البياناتDM في دعم المحتوى المعلوماتي للقوائم المالية المستقبلية، مما يؤدى إلى تحسين مستوى جودة القوائم المالية. وذلك من خلال استخدام أسلوب التنقيب في البياناتDM كتقنية جديدة والاستفادة منه في دعم المحتوى المعلوماتي للقوائم المالية المستقبلية، وزيادة حجم الاستثمارات من خلال تعزيز كفاءة قرارات الاستثمار، بالإضافة إلى تطوير إطار محاسبي لاختبار الدور الوسيط لقرارات الاستثمار في العلاقة بين المحتوى المعلوماتي للقوائم المالية المستقبلية وكفاءة السوق المالي المصري .

توصلت الدراسة إلى العديد من النتائج أهمها: أن إعداد القوائم المالية المستقبلية يعتمد بشكل كبير على سلامة التنبؤات والتقديرات المالية، واستخدام أساليب علمية متطورة في عملية الإعداد، لكي تعكس القوائم المالية المستقبلية صورة أقرب ما يكون إلى الواقع. ونظراً لأن البيانات المحاسبية تتسم بالتعقيد مما يتطلب استخدام منهج علمي متطور لتحليل هذه البيانات، تتوافر فيه مجموعة من الخصائص من أهمها: القدرة التنبؤية العالية مع مراعاة طبيعة البيانات، والتعامل مع المتغيرات المالية وغير المالية، وهو ما توفره أدوات التنقيب في البيانات.

توصلت الدراسة إلى ضرورة قيام منشآت الأعمال المصرية بإعادة النظر في النموذج التقليدي الحالي لإعداد القوائم المالية المستقبلية، بحيث تعكس وبشكل أكبر العمليات المستقبلية للشركة، والافتراضات التي تبنى عليها لدعم ثقة المستثمرين في تلك القوائم، وما تتضمنه من معلومات محاسبية ومالية، والذي ينعكس بالتالي على تعزيز القرارات الاستثمارية. كذلك ضرورة الإفصاح عن المعلومات المالية المستقبلية، ويفضل أن تكون في شكل قوائم مالية مستقبلية يتم إعدادها بشكل دوري، وفى حالة تعذر ذلك فإنه يجب على تلك الشركات الإفصاح عن بعض البنود أو العناصر ذات التأثير في قرارات الاستثمار.

التعليق على الدراسات السابقة:

ترى الباحثة أن العديد من الدراسات السابقة التي تناولت تقنية التنقيب في البياناتDM تناولتها من وجهة النظر الاقتصادية أو الإدارية، أو من خلال إجراء مسح شامل للتعرف على مهام وأهداف التنقيب في البيانات DM ، بينما ركزت الدراسات المحاسبية التي تناولت التنقيب في البياناتDM على استخدام أحد أساليبها للتنبؤ بالمخاطر الائتمانية للبنوك بصفة عامة، ومنح القروض بصفة خاصة، وبعض الدراسات تناولت أثر استخدام التنقيب في البياناتDM على جودة التقارير المالية، أو قدرتها على رفع كفاءة المراجعة، ودعم مسؤولية مراجع الحسابات عن كشف الاحتيال والخطأ، ومدى الحد من المعلومات المضللة في التقارير المالية.

تسعى الباحثة إلى تقديم دراسة حالة عن استخدام الأساليب المختلفة للتنقيب في البيانات ، لتحديد أفضلها وأكثرها دقة كي يتم اعتماد للتنبؤ بقدرة العميل على سداد القرض وفوائده، لمساعدة إدارة البنوك للحد من المخاطر الائتمانية التي قد يواجهها البنك، وبالتالي زيادة جودة القوائم المالية من خلال الحد من المخاطر الائتمانية. نظراً لأن أغلبت الدراسات العربية والمصرية بصفة خاصة اعتمدت على تحليل لقوائم استبيان، ولكن هذا البحث يقوم على

دراسة حالة باستخدام برنامج Anachoda على مجموعتين من بيانات أحد البنوك أحدها يتم استخدامها للاختبار والمجموعة الأخرى هي مجموعة بيانات ناتجة عن تجارب البنك السابقة في منح القروض وقدرة العملاء على السداد من عدمه وفقا لمجموعة المتغيرات أو الخصائص التي يصدر البنك اعتماداً عليها قراراته الائتمانية.

٢/٢ التنقيب في البيانات Data Mining (DM) كأحد أدوات الذكاء الاصطناعي:

ظهر مصطلح التنقيب في البيانات في الولايات المتحدة في منتصف الستينيات، وهو مصطلح يجمع بين عدة تعريفات للإحصاءات وتكنولوجيا المعلومات وقاعدة البيانات والذكاء الاصطناعي ((Artificial Intelligent(AI) والتعلم الآلي ومفاهيم البيانات.

١/٢/٢ مفهوم التنقيب في البيانات (DM):

وقد ظهرت عدة تعاريف لمفهوم تنقيب البيانات DM . عرفها (العلى وآخرون، ٢٠٠٦) بأنها عبارة عن تحليلات لكمية كبيرة من البيانات لإيجاد قواعد وأمثلة ونماذج يمكن أن تستخدم لتقود وتدل صاحب القرار، وتتنبأ بالسلوك المستقبلي، وعرفها (Al-Azmi,2013) بأنها عملية تحليل قواعد البيانات الكبيرة، عادة ما تكون مستودعات البيانات أو الإنترنت لاكتشاف معلومات جديدة وأنماط وسلوكيات مخفية أي إنها عملية آلية لتحليل كميات هائلة من البيانات لاكتشاف السمات والأنماط المخفية والتنبؤ بالاتجاهات المستقبلية والتنبؤ بالفرص المحتملة .

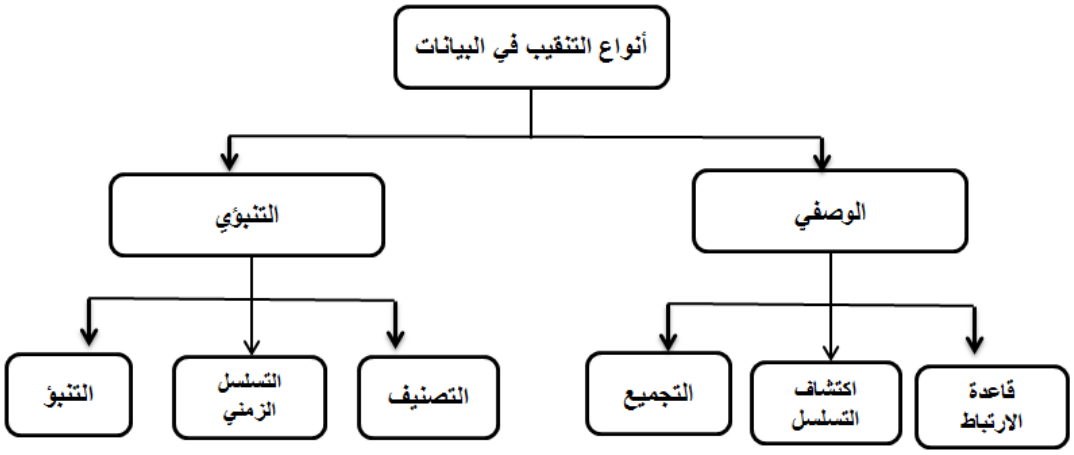
من أهم مميزات التنقيب في البيانات DM أنه يمكن إجراءه على أي نوع من أنواع البيانات (سواء كانت نصية أو رقمية) طالما كانت البيانات مفيدة للتطبيق المستهدف، ويمكن اعتبار التنقيب في البياناتDM تقييماً لتكنولوجيا المعلومات وتضافراً بين عدة تخصصات ومجالات تطبيقية ذات صلة.

فالهدف الرئيسي للتنقيب في البيانات الوصفية هو فهم البيانات (ما حدث)، ثم التنبؤ من خلال استخدام الماضي لفهم المستقبل(ما يمكن أن يحدث)، ثم الوصول إلى أفضل نتيجة (ما يجب أن يحدث) وبالتالي اتخاذ قرارات مستقبلية، حيث تستخدم تقنيات التصنيف والتوصيف وتوحيد وتصوير البيانات لتحويلها إلى معلومات مفيدة لأغراض تحسين البيانات وفهم الأعمال.

مما سبق عرفت الباحثة "التنقيب في البيانات DM" بأنه عملية لاستخراج المعرفة من البيانات الموجودة من خلال استخراج معلومات مخفية وغير معروفة وصالحة وقابلة للتنفيذ من قواعد بيانات كبيرة، ثم استخدام هذه المعلومات لاتخاذ قرارات هامة، كالنظر في إدارة العلاقة مع العملاء، ومحاولة تحليل الحجم الهائل لبيانات العملاء، وتصنيفها وفقا لسلوك العميل وتوقعاته، وتحليلات السوق.

ويمكن توضيح أنواع التنقيب في البيانات من خلال الشكل التالي:

شكل رقم (١) أنواع تنقيب البيانات



المصدر : (البدوي ، ٢٠١٧)

٢/٢/٢ أساليب وأدوات التنقيب في البيانات DM :

وتتمثل أساليب وأدوات التنقيب في البيانات DM(البدوي، ٢٠١٧)، فيما يلي:

شجرة القراراتDecision Trees، الغابات العشوائية Random Forest، ونموذج الانحدار اللوجستيLogistic Regression، والشبكات العصبية Neural Networks، والخوارزميات الجينية Genetic Algorithms، السلاسل الزمنية، والمنطق الغامض Fuzzy logic، التحليل التمييزي Discriminant Analysis، الاسناد Boosting، المجاور الأقرب-K-Nearest Neighbor(k NN) .

تتمثل معايير تقييم اختيار الأسلوب المناسب للتنقيب عن البيانات (Sharma, A. et al., 2012)، فيما يلي:

- الأداء حيث يركز على قدرة الأداء وسهولة معالجة البيانات تحت ظروف مختلفة.
- القدرة على الاستخدام.
- امكانيات الأداة حيث يتضمن مجموعة من الامكانيات والتقنيات والمنهجيات للتنقيب عن البيانات.
- الدعم حيث أن المقصود به أداء العديد من الوظائف الثانوية المطلوبة في عملية التنقيب في البيانات وتتضمن هذه المهام اختيار البيانات ، وتنظيفها ، وفلترة البيانات.
- الكفاءة في إظهار النتائج.
- نظام التشغيل هل البرنامج يعمل على نظم تشغيل متنوعة أم نظام تشغيل محدد.

جدير بالذكر أن هناك العديد من المنتجات والأدوات والبرامج التي تقوم بعملية تنقيب البيانات والتي تنتج عن شركات كبيرة مثل (SPSS, IBM, Microsoft, Anaconda, Panda) وغيرها من الشركات المعروفة وتستخدم معظم أدوات التنقيب لغة واحدة للعرض.

٣/٢/٢ تصنيفات منهج تنقيب البيانات:

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 التي تهدف إلى التنبؤ، بمقارنة السجلات الشبيهة بالسجل المراد التنبؤ لها، وتقدير القيمة المجهولة لهذا السجل بناء على معلومات تلك السجلات والشبكات العصبية.

٢) الانحدار Regression: 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: 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)التجميعClustering: بمعنى وصف الخصائص العامة للنماذج وتنظيم البيانات في مجموعات

٦)المرئية Visualization: مرئية النتائج يساعد في تسهيل ملاحظات، مخرجات خوارزميات تنقيب البيانات وفهمها.

٤/٢/٢ مجالات الاستفادة من التنقيب في البيانات في بيئة الأعمال:

امتدت الاستفادة من تطبيقات تنقيب البيانات إلى عدة مجالات، وهي:

١-الأعمال المصرفية: كتحليل مخاطر القروض، تحديد العملاء الذين يمتازون بالولاء، تحديد نماذج مزوري استخدام بطاقة الائتمان.

٢-الاحتيال المالي: تحديد الغش في تبادل الأسهم والقوائم المالية ، ايجاد الارتباط الخفي بين المؤشرات المالية المختلفة ، تحديد خصائص الصناعة والسوق والتنبؤ بأداء الفرد والأسهم، تحديد قواعد تجارة الأسهم من بيانات السوق التاريخية (Ramachandran, 2011).

٣-التسويق: إيجاد العلاقة بين الخصائص الديموغرافية للزبون، تحليل سلة السوق، تجزئة السوق، تحديد الغش في الصفقات، تحديد نماذج الشراء للزبائن، تحليل الاتجاه (بمعنى اكتشاف الزبون المثالي لهذا الشهر مقارنة بالشهر الماضي) (Yigitbasioglu et al.,) (Xiao et al., 2010, 2012

٤-ادارة علاقات العملاء: اتخاذ الاجراءات اللازمة للاحتفاظ بالعملاء والزبائن، والسعي نحو اكتساب عملاء جدد.

٥- التجارة الالكترونية.

٦- تحليل الضمان.

٧-رقابة الجودة أو تحليل الخطأ: تحديد الخصائص المحيطة بالمنتجات المتضررة، ومن ثم تحسين الجودة من خلال اجراء التغيرات.

٨- الاستخدام: تحديد خصائص الاستخدام، تحديد خصائص أفراد الأداء العالي كالتعليم وسنوات الخبرة والمزايا الشخصية، والمهارات لتأسيس هيكل الاستخدام للمساعدة في استقطاب واستخدام الأفراد ذوى الأداء الأفضل.

٩- التعليم لاستقطاب وابقاء الطلاب

١٠- تحليل قائمة الخدمة في المطاعم

وترى الباحثة استناداً على ما سبق أن تنقيب البيانات هو عملية اكتشاف معلومات قيمة من مخازن البيانات الكبيرة للإجابة على أسئلة الأعمال الهامة، كما إنه يكشف النقاب عن العلاقات والاتجاهات والأنماط والاستثناءات الضمنية والشاذة التي كانت مخفية في التحليل البشري. ويمكن استخدامه بطرق مختلفة، مثل: تسويق قواعد البيانات، وإدارة مخاطر الائتمان، وكشف الاحتيال، وتصفية البريد الإلكتروني، أو حتى تمييز مشاعر العملاء أو أرائهم.

تنقيب البيانات يستخدم الوصف أساساً لوصف الأنماط والاتجاهات فى البيانات، بينما يتم تقدير قيمة المتغير الرقمي المعتمد على تقدير واحد أو أكثر من المتغيرات المستقلة، وتستخدم تقنية التصنيف العثور على قيمة المتغير التابع بناء على بعض المتغيرات المعالجة، ولكن فى هذه الحالة يكون المتغير التابع قاطعاً وليس ذا قيمة عددية، ويعد اكتشاف الغش مثالاً شائعاً على مشكلة التصنيف، ويمكن استخدام التنبؤ فى تقديرات القيم المستقبلية.

٣/٢ المخاطر التي يواجهها القطاع المصرفي:

تخضع البنوك بصفة عامة للعديد من المخاطر منها المتعلق بالهيكل التكنولوجي والمالي، ومنها ما قد يؤثر على سمعتها ، وبعضها مستمد من البيئة التنظيمية والاجتماعية. وتداخل هذه المخاطر بشكل متبادل ولديها بعض التقاطعات التي تجعل من الصعب عزلها وتحديدها.

على الرغم من أن البنوك الإلكترونية تقدم العديد من الخدمات والمزايا المالية الالكترونية للعملاء من حيث الراحة وسهولة الاستخدام والتكلفة المنخفضة وما إلى ذلك، إلا أنها تحمل في طياتها العديد من المخاطر التقليدية والمكتسبة (Embark,2022):

أ) مخاطر التشغيل/المعاملات Operational /Transactional risks : هي أكثر أنواع المخاطر المصرفية الإلكترونية شيوعًا. تتضمن المعالجة غير صحيحة للمعاملات، ومساومات في سلامة البيانات، وخصوصية البيانات، والسرية، واختراق أنظمة البنك، وعدم قابلية تنفيذ العقود، وما إلى ذلك. والناتج عن الأخطاء التكنولوجية، وعوامل الإهمال البشري(Zarei, 2011).

ب)المخاطر الاستراتيجية Strategic risks : هي المخاطر التي تؤثر على أرباح الاستثمار على المدى الطويل بسبب قرارات العمل غير الملائمة، والأداء غير الملائم.

ج)مخاطر التكنولوجيا Technology risks: نشأت المخاطر التكنولوجيا الحالية والمحتملة نتيجة للخسائر والانتهاكات والتشوهات وتعطل أجهزة الكمبيوتر والبرامج والخدمات عبر الإنترنت ، إلخ(Mogos, et al., 2021).

د) مخاطر الامتثال: هي المخاطر الناتجة عن عدم الامتثال لتشريعات الدولة والمبادئ الأخلاقية، يجب أن تكون المؤسسة المالية في وقت لاحق على دراية بالقانون الذي يمكن الحصول عليه والتأكد من تطبيقه على كافة الفروع. تشمل مخاطر الامتثال أيضًا الحاجة إلى الحفاظ على سرية بيانات المستهلكين (Mogos, et al., 2021).

هـ)المخاطر الأمنية الإلكترونية: تعتبر الخطر الرئيسي للخدمات المصرفية الإلكترونية، حيث تتقاطع بعض المشاكل المحددة مع فئات المخاطر مثل خرق الأمن والسماح بالوصول غير المصرح به إلى معلومات العملاء. مثل هذه القضايا الأمنية قد تعرض البنك لمخاطر قانونية ومخاطر تتعلق بالسمعة (Mogos, et al., 2021).

و) مخاطر السمعة: ناتجة عن عدم قدرة البنك على أداء الإجراءات والوظائف التي يتوقعها العميل، أو فشل النظام في الخروقات الأمنية والوصول إلى الحسابات، مما قد يتسبب في فقدان الثقة ومغادرة البنك. يمكن أن يتضرر هؤلاء بسبب السلوك غير الأخلاقي والإجرامي للأفراد والجماعات لأنه يمكن اعتباره تهديدًا

لنظام البنوك الإلكترونية ككل، ولكن أيضًا للمستهلكين (Mogos, et al., 2021).

ز) مخاطر الائتمان: مخاطر تتعلق بعدم قدرة العميل على الوفاء بالتزاماته المالية. حيث تقدم البنوك الإلكترونية مزايا للعملاء بمنح قرض من أي مكان في العالم، ولكن هذا يزيد من صعوبة قدرة البنك على تقييم الجدارة الائتمانية للعميل(Safari, et al., 2020)

ح)مخاطر أسعار الفائدة: المتعلقة بتحركات أسعار الفائدة. تتيح البنوك الإلكترونية للعملاء فرصة مقارنة أسعار الفائدة المختلفة، وتحتاج البنوك إلى التحرك بسرعة للاحتفاظ بالقروض والودائع من العملاء أو الحصول عليها.

ط)مخاطر الصرف الأجنبي: قد تشجع الخدمات المصرفية الإلكترونية المقيمين في دولة واحدة على إجراء معاملات بعملتهم المحلية، ونتيجة لذلك قد يقوم العملاء بالعديد من إجراءات المضاربة (Ahmad, et al., 2011).

وسوف تقتصر الباحثة خلال هذا البحث على دراسة مخاطر الائتمان، وهي جزءاً لا يستهان به من العمل المصرفي خصوصاً مع ارتفاع حدة المنافسة وزيادة حجم المعاملات المصرفية، وبالتالي فإن حسن إدارة هذه المخاطر المحتملة تعتبر من العوامل التي تساعد البنك على نجاحه وضمان استمراره في السوق المصرفية بعائد مرضي ومخاطر متدنية.

٤/٢ المخاطر الائتمانية وانعكاساتها على جودة القوائم المالية للبنوك:

تعد مخاطر الائتمان من أهم المخاطر التي تواجه البنوك لما لها من آثار سلبية على ربحية المؤسسات المالية، وقيمتها السوقية، والتزاماتها، وحقوق المساهمين، وبالتالي التأثير على جودة القوائم والتقارير المالية.

عرف(البنك المركزي، ٢٠١٨) الائتمان بأنه: "الثقة التي يوليها البنك لشخص ما سواء أكان طبيعياً أم معنوياً"، بأن يمنحه مبلغ من المال لاستخدامه في غرض محدد، أو تقرير حدود يضعها تحت تصرفه لفترة محددة متفق عليها، وبشروط معينة مقابل عائد مادي متفق عليه، يتمثل في الفوائد والعمولات وبضمانات تُمكن البنك من استرداد قرضه في حال توقف العميل عن السداد وفقا للسياسة الائتمانية للبنك.

حددت لجنة بازل للإشراف المصرفي مخاطر الائتمان ، "بأنها إمكانية الخسارة الجزئية أو الكلية للقرض الحالي بسبب عدم سداد هذا القرض في الوقت المتفق عليه أي في الوقت المحدد، وبالتالي فإن الزيادة في الائتمان ستؤدي المخاطر إلى زيادة التكلفة الحدية للديون وحقوق الملكية. (Ekinci, Poyraz, 2019)

وتمثل مخاطر الائتمان درجة تذبذب العائد الفعلي لعملية الإقراض من العائد التعاقدي، نتيجة لأسباب داخلية تتمثل في ضعف إدارة البنك، أو لأسباب خارجية تتمثل في عدم القدرة أو التأخر في السداد، أو لأسباب اقتصادية (نجيب وآخرون، ٢٠٢١).

وفقا للتعريفات السابقة، ترى الباحثة أن مخاطر الائتمان هي الخسائر المالية المحتملة الناتجة عن عدم قدرة العميل على الوفاء بالتزاماته في المواعيد المحددة نتيجة وجود ضائقة مالية يتعرض لها المقترض وتحد من قدرته على السداد، سواء بالنسبة لأصل القرض أو الفائدة، أو كليهما. وهناك عدة أسباب لعدم قدرة المقترض على السداد:

أ- وجود خلل في تقدير الجدارة الائتمانية والقوة المالية للمقترضين، أو لم يتم تحديده بشكل كافٍ.

ب- افتقار موظفي البنك لمهارات معالجة طلبات القروض.

ج- عدم وجود نظام مناسب للتصنيف وقياس وإدارة مخاطر الائتمان عبر المواقع الجغرافية وخطوط الإنتاج ، وعدم موثوقية البيانات المستخدمة لإدارة الائتمان والمخاطر المرتبطة بالإقراض.

د- سوء النية من جانب المقترض، أي سوء النية من أجل الحصول على التسهيلات الائتمانية بسرعة، وهو سلوك متعمد من جانب المقترض.

هـ- اهتمام البنوك بتوسيع الائتمان من أجل زيادة أرباحها بغض النظر عن درجة المخاطر التي تتعرض لها.

و- عدم المتابعة المستمرة لأوضاع المقترضين والمشاريع الممولة.

ز- ظروف طارئة غير متوقعة تؤدي إلى تدهور الوضع الاقتصادي في البلاد

استناداً على ما سبق تؤكد الباحثة أنه في حالة تقييم الجدارة الائتمانية للعميل يجب أن تنظر البنوك إلى العناصر الخمسة التالية لتقييم مخاطر الائتمان، وهي: تاريخ الائتمان، والقدرة على السداد، ورأس مال المقترض، وشروط القرض، والضمانات المرتبطة بها.

كلما زادت المخاطر الائتمانية كلما كان لها أثر سلبي على ثقة المستثمرين ومتخذ القرار على الثقة في المعلومات المحاسبية والبيانات المالية في القوائم المالية، حيث يتم احتساب سعر السهم والعائد على حقوق الملكية و معدل الربحية والعائد على حقوق المساهمين من خلال المعلومات المحاسبية المعلنة في القوائم المالية، وبالتالي فمن الأهمية ضرورة السعي نحو الحد من المخاطر الائتمانية من خلال استخدام التقنيات التكنولوجية الحديثة، لما له من تأثير على جودة القوائم المالية وانعكاساتها على التقارير المالية للبنوك (Pere, 2010, Digalaki, 2019).

وتؤكد الباحثة أن البنوك تحاول الحد من مخاطر الائتمان من خلال التأكيد على مجموعة من الملامح الأساسية لمنح الائتمان والتي تتمثل فى: الثقة وتعكس درجة المخاطرة المقبولة، ومبلغ الائتمان ويمثل حجم الموارد القابلة للتوظيف لدي البنك ومدي ملاءة العميل وقدرته علي السداد، والغرض من الائتمان والذي يتمثل في تمويل الأنشطة الاستثمارية أو الجارية، فترة الائتمان وتحدد بفترة السداد على دفعات أو دفعة واحدة والضمانات التي تضمن للبنك استرداد القروض.

تعزو أهمية مخاطر الائتمان في معظم عمليات البنوك إلى تأثيرها على القوائم المالية للبنك، إذ أن معظم الأصول المكونة لمحافظ البنوك تتعرض للخطر، نتيجة عدم قدرة العميل على سداد قيمة القرض وفوائده عند تاريخ الاستحقاق المحدد في شروط العقد الائتماني. وبالتالي ينعكس الأثر على بنود داخل الميزانية مثل: القروض والسندات، وبنود خارج الميزانية: مثل خطابات الضمان والاعتمادات المستندية(حسين،٢٠٢١). وفي هذا السياق أشار (الجابري، ٢٠١٧) الى أن المخاطر الائتمانية هي من المخاطر المنتظمة العامة التي لا يمكن تجنبها أو إلغاؤها، ولكن يمكن التعايش معها والتقليل من آثارها غير المرغوبة، لما لها من تأثير على صافي دخل البنك والقيمة السوقية لحقوق المساهمين وبالتالي قيمة البنك، ومن أجل ذلك أصبحت البنوك في الوقت الحاضر تولي أهمية قصوى للإفصاح عن تلك المخاطر لما لها من تأثير على جودة وعدالة القوائم والتقارير المالية.

تزايد معدلات القروض المتعثرة، أو عدم وجود ضمانات كافية يؤثر سلبًا على جودة القوائم المالية للبنوك. فهناك علاقة سلبية بين القروض المتعثرة وكفاية رأس المال من ناحية والربحية مقاسة بمعدل العائد على حقوق الملكية من ناحية أخرى. وعلى النقيض، فإن نسبة كفاية رأس المال لها علاقة إيجابية بالأداء المالي للبنك والذي ينعكس على جودة القوائم المالية، فإذا كان رأس مال البنك قوي فإنه يصبح قادراً على امتصاص خسائر القروض المحتملة، وتجنب الفشل المالي والإفلاس. وهو ما يؤكد على وجود علاقة سلبية بين القروض المتعثرة والبنك(Hassoon, et al., 2022).

كما أن الزيادة في مخاطر الائتمان المصرفي تؤدي إلى مشاكل تدريجية في السيولة والملائمة المالية. وهو ما يمثل تهديداً خطيراً على جودة القوائم المالية للبنوك، لتأثيره على قدرة العملاء على قراءة وتقييم القوائم المالية لتحديد أوجه استثماراتهم، وتحديد البنوك الموثوق بها للاستثمار(Samuel, 2015).

وترى الباحثة أن مما يؤكد على تأثير مخاطر الائتمان على جودة القوائم المالية، إصدار لجنة بازل للرقابة المصرفية في ديسمبر ٢٠١٥، إرشادات رقابية توضح كيفية ممارسة مخاطر الائتمان المتمثلة في خسائر الائتمان المتوقعة ومعرفة التفاعل بين النهج المحاسبي الجديد وجميع متطلبات إطار بازل (شاهين والبغدادي ، ٢٠١٩)، وكذلك قيام البنك المركزي المصري بوضع إطار لتصنيف الأصول المعرضة لخطر الائتمان في مجموعات وفقا لقواعد وشروط تفصيلية، تعتمد على معلومات عن العميل، ونشاطه ووضعه المالي، ومدي انتظامه في السداد. ويقوم البنك بحساب مخصصات لاضمحلال الأصول المعرضة للخطر على أساس نسب محددة من قبل البنك المركزي المصري لأغراض إعداد القوائم المالية وفقا لمعايير المحاسبة المصرية، ويتم تجنب احتياطي المخاطر البنكية العامة ضمن حقوق الملكية خصماً على الأرباح المحتجزة بمقدار تلك الزيادة، ويتم تعديل الاحتياطي دورياً بالزيادة أو النقص بحيث يعادل دائماً مبلغ الزيادة بين المخصصين ويعد هذا الاحتياطي غير قابل للتوزيع (قواعد إعداد وتصوير القوائم المالية للبنوك وأسس الاعتراف والقياس المعتمدة من مجلس إدارة البنك المركزي، ٢٠٠٨).

مما سبق ترى الباحثة أن الائتمان المصرفي يمثل مؤشر للمخاطر الائتمانية وفقاً لمجموعة من العوامل وهي: أجله، طبيعة المقترض ونشاطه، ونشاط الاقتراض، والضمانات المقدمة للاقتراض.

ترى الباحثة أنه لا يجب الاقتصار عند تحليل المخاطر الائتمانية على الأساليب التقليدية لتحليل النسب، بل يجب التوجه نحو الأساليب التكنولوجية الحديثة، لتحديد وقياس ومراقبة مخاطر الائتمان، وهو ما دعى الباحثة إلى دراسة أثر استخدام التنقيب في البيانات DM باعتباره أحد أدوات الذكاء الاصطناعي المستحدثة.

٥/٢ دور التنقيب في البياناتDM في الحد من المخاطر الائتمانية في القطاع المصرفي، مما يؤثر على جودة القوائم المالية:

تتمثل أهمية التنقيب في البيانات DM في القطاع المصرفي على تحقيقه لثلاثة أهداف رئيسية: الوصف والتنبؤ والتصور.

- يُمكن التنقيب في البيانات DM من استكشاف المعارف القيمة والأنماط المثيرة للاهتمام المخبأة في الكم الهائل من البيانات

- يساعد القطاع المصرفي على استخدام جمع البيانات في عمليات صنع القرار التي يقوم بها في مجالات: التسويق، وإدارة مخاطر الائتمان، والكشف عن غسيل الأموال، وإدارة السيولة، والمصارف الاستثمارية، والكشف عن المعاملات الاحتيالية (Tiwari, 2010). ففي حالات تعثر السداد في هذه المجالات يمكن أن يؤدي إلى نتائج غير سارة ، مثل: خسارة العملاء أمام المنافسة، والخسارة المالية، وفقدان السمعة، والغرامات الباهظة من الجهات التنظيمية.

- يُمكن التنقيب في البيانات DM موظفي القطاع المصرفي من تحديد الأنماط والاتجاهات في البيانات واكتشاف المشكلات أو مجالات الفرص(Ngai, et. al., 2009).

- يستخدم التنقيب في البيانات DM في التنبؤ بإجمالي الربح والتدفقات النقدية، والأداء المالي وتحليل النسب المالية من خلال الاعتماد على الشبكات العصبية، وشجرة القرارات وشبكة بايزن، وأساليب التصنيف والتنبؤ والارتباط والانحدار وغيرها من أساليب التنقيب والتي أعطت مرونة للتعامل مع المتغيرات المالية التي تمثل جزءاً مهماً في عملية التحليل المالي والتنبؤ بالأرباح والتدفقات النقدية. (Amani, et al., 2017)

- يساعد استخدام خوارزمية Cartفي شجرة القرارات في تحليل المؤشرات المالية للبنوك، مما يساعد على التنبؤ بسلوك العملاء المتعاملين مع البنك في المستقبل، من خلال تصنيف

العملاء والنمط الجامع بينهم في مختلف المجالات والخدمات المقدمة من البنوك(على، ٢٠١٨).

- **الاستمرار في نظام التنقيب التنبؤي في ظل نظام واضح ومتطور ، يحافظ على السرعة مصحوباً بالدقة في إنجاز الربح المؤكد.**

- **استخدام التنقيب التنبؤي في النظام المصرفي ينبه عميل بطاقة الائتمان إلى احتمال وجود رسوم احتيالية.**

- **يؤكد التنقيب في البياناتDM في القطاع المصرفي، على استخدام نظام لدعم اتخاذ القرارات يستند إلى تقنيات استخراج البيانات لتحسين نوعية عملية الإقراض في المصارف (Ionita & Ionita, 2011). إن إجراء التحليل اليدوي يمكن أن يشوبه بعض أوجه القصور لأن كميات البيانات التي يمكن تحليلها يدوياً محدودة، وبالتالي فإن القرارات قد لا تكون دقيقة كما كان مستهدفاً. على سبيل المثال: قد يكون من الممكن دفع أقساط القروض بانتظام رغم وجود اتجاه سلبي مقلق في معدل دوران العملاء، وقد يكون حساب العميل على وشك التخلف عن السداد، وليس من السهل اكتشاف هذه الارتباطات من خلال العمليات اليدوية. ويفترض أن المعلومات القيمة مخبأة في هذا الحجم من البيانات التشغيلية والتاريخية التي يمكن استخدامها في عملية اتخاذ القرارات الحاسمة إذا تم اكتشافها واستخدامها بواسطة أدوات فعالة (Kazi& Ahmed, 2012).**

- **مع تصاعد حدة المخاطر الائتمانية في المصارف في بيئة السوق التنافسية للغاية، والتي قد تفسد اختيارات العملاء، تحتاج المصارف إلى أن تكون استباقية في تحليل تفضيلات العملاء، وملفات التعريف الخاصة بهم، وأن تضبط منتجاتها وخدماتها وفقا لذلك للحفاظ على قاعدة العملاء (Bhambri, 2011) فمن خلال تقسيم العملاء إلى عملاء سيئين وعملاء جيدين، يستطيع البنك أن يخفض الخسائر قبل فوات الأوان (Kazi & Ahmed, 2012)، ومن خلال تحليل أنماط المعاملات يمكن للمصرف أن يتتبع معاملات الاحتيال قبل أن يؤثر على أرباحه، وهذه قدرات مرغوبة للغاية.**

- **يساعد التنقيب التنبؤي صانع القرار نحو الاختيار الأصوب، كما يستخدم التنقيب الوصفي تقنيات التحسين لتحديد أفضل البدائل لتقليل أو تعظيم بعض الوظائف الموضوعية، يمكن أن**

تكون التقنيات الرياضية والإحصائية للتنقيب عن البيانات التنبؤية جنبًا إلى جنب مع التحسين لاتخاذ القرارات، التي تأخذ في الاعتبار عدم اليقين في البيانات، سواء كان الهدف هو التقرير بأثر رجعي أو مستقبلي(صالح، ٢٠١٦).

- بالإضافة إلى وجود مجالات مختلفة يُمكن فيها استخدام التنقيب في البياناتDM في القطاعات المصرفي، مثل: تقسيم العملاء والربحية، وتحليل الائتمان، وتوقع الخطأ في الدفع والتسويق والمعاملات الاحتيالية، وترتيب الاستثمارات، وتحسين محافظ الأسهم، وإدارة النقد وعمليات التنبؤ، وتقييم مقدمي طلبات القروض عالية المخاطر، وعملاء بطاقات الائتمان الأكثر ربحية والبيع المتبادل.

- يساعد تنقيب البيانات في القطاع المصرفي (على، ٢٠١٨) على تقليل نسبة التكاليف والوقت والجهد حول اعتبار ما سيكون عليه المستقبل في المجال المصرفي بصورة خاصة.

- يساعد على اكتشاف علاقات جديدة من واقع البيانات المخزنة واستغلال البيانات بصورة جيدة.

- التنقيب في البيانات يساعد مسئولي القطاع المصرفي على تحليل الانحرافات في سداد الديون أو التلاعب في الحسابات المالية.

- التنقيب في البيانات يساعد على عمليات التقدير المالي للمعاملات التجارية، عن طريق تحليل البيانات المجمعة والتنبؤ بالمعلومات المستقبلية، وهو ما يمكن استخدامه فى مجال التخطيط حيث تقدم معلومات عن التنبؤ بحجم الطلب من العملاء على المدى القصير، كما

٦/٢ أساليب التنقيب في البياناتDM المستخدمة في التنبؤ بمنح القروض في القطاع المصرفي:

تستخدم تقنية التصنيف في القطاع المصرفي للكشف عن الاحتيال (الاحتيال في الشركات والائتمان على السواء) (Ngai et al, 2011) وتبنى هذه النماذج عادة باستخدام نموذج شجرة القرارات أو نموذج الانحدار اللوجستي ، نموذج الغابات العشوائية، نموذج الشبكة العصبية.

مما يؤكد على ضرورة التوجه نحو زيادة تطبيقات الذكاء الاصطناعي داخل البنوك ما توصلت اليه دراسة (Digalaki,.2019) أن حجم الانخفاض في التكلفة الاجمالية المحتملة للبنوك سيبلغ ٤٤٧ مليار بحلول عام٢٠٢٣.

١) شجرة القرارات:

هي تقنية تعلم خاضعة للإشراف يمكن استخدامها لكل من مشاكل التصنيف والانحدار، لكنها في الغالب مفضلة لحل مشاكل التصنيف، لإنه مصنف منظم على شكل شجرة، حيث تمثل العقد الداخلية ميزات مجموعة البيانات، وتمثل الفروع قواعد القرار، وتمثل كل عقدة طرفية نتيجة شجرة القرارات، حيث توجد عقدتان هما: عقدة قرار وعقدة طرفية .

تُستخدم عقد القرار لاتخاذ القرار ولها فروع متعددة، في حين أن العقد الطرفية هي نتاج هذه القرارات ولا تحتوي على أي فروع أخرى. شجرة القرار هي رسم بياني على شكل شجرة يمثل العلاقات بين مجموعة من المتغيرات، تُستخدم لحل مشاكل التصنيف والتنبؤ، حيث يتم تصنيف الحالات إلى فئة من فئتين، عادةً ما تكون إيجابية وسلبية، أو متعثرة وغير متعثرة في حالة تصنيف قروض البنوك. يتم تمثيل هذه النماذج وتقييمها بطريقة من أعلى إلى أسفل. يتضمن تطوير أشجار القرار مرحلتين: بناء الاشجار وتقليم الاشجار(Embarh, et al., 2022).

وتستخدم شجرة القرارات للتنبؤ بالقيمة، فعلى سبيل المثال، بدلاً من تصنيف طلبات القروض الجديدة، تحاول التنبؤ بالمبالغ الافتراضية المتوقعة مستقبلاً لطلبات القروض الجديدة. بمعنى آخر: القيم المتوقعة هي قيم رقمية، وبالتالي فهي تتطلب تقنيات النمذجة التي يمكن أن تأخذ البيانات الرقمية كمتغيرات مستهدفة (أو متوقعة).

تستخدم شجرة القرارات لمقارنة الأداء المالي للبدائل المتوقعة لاتخاذ القرارات. كما تستخدم في تقييم البدائل المختلفة لتوسيع الطاقة الإنتاجية حين يكون الطلب غير مؤكد والقرارات متعاقبة، وغالبا ما يستخدمها متخذ القرار للحصول على وصف أو صورة واضحة لبدائل القرار ونتائجها المحتملة.

قامت دراسة (Mandalaa, et al, 2012) باستخدام شجرة القرارات كأحد أدوات تنقيب البيانات، لدراسة تقييم جدوى طلب القروض، لتقييم ما إذا كان العميل لديه أنشطة تجارية ممكنة وقابلة للتسويق ومربحة ويمكن سدادها على أساس الوقت، وذلك وفقا لدراسة معايير الأداء

المالي للشركات وهي رأس المال، السمعة الشخصية، القدرة على سداد القرض شهرياً، والضمانات التي يقدمها.

تستخدم شجرة القرار، من خلال خوارزمية التصنيف والانحدار للاحتفاظ بالعملاء في القطاع المصرفي والتعرف المستقبلي عن توجهاتهم، والذي يعتمد أيضًا على طريقتين: طرق التصنيف وطرق التنبؤ بالقيمة.

أشار (Embarh, et al., 2022) بعد استخدام شجرة القرار، كلما انخفض عامل الثقة، تم إجراء المزيد من التقليم(أي استبعاد بعض الفروع من الشجرة لتحقيق الدقة). وقد استخدمت الدراسة عوامل ثقة مختلفة وتوصلت إلى أن زيادة عامل الثقة تكون الدقة أعلى، ومع عامل ثقة ٠.١٥ فإن أفضل دقة هي ٦٢.١٢٪ ومع عامل ثقة ٠.٢٥ فهي ٦٣.٣٩٪. وهذا يعني أنه عندما يتم التقليم بشكل أقل ، تتحسن الدقة. ومع ذلك، فمن المعروف أنه كلما تم التقليم بشكل أقل، كلما أصبح النموذج أكثر تجهيزًا ويمكن أن يتسبب في نتائج سيئة أثناء التصنيف. وبالتالي تتحسن الدقة عند تقليل التقليم (عندما يصبح النموذج أكثر من اللازم).

توصل (على، ٢٠١٨) أن خوارزمية CART في شجرة القرار تتعامل بسهولة مع القيم المتطرفة، فهذه القيم تؤثر سلباً على قيم بعض النتائج الإحصائية مثل تحليل العناصر والانحدار الخطي، ولكن التقسيم في شجرة القرار هنا يسهل التعامل مع القيم المتطرفة فيقوم بعزلها ووضعها في عقدة منفصلة.

كما تتسم CART بالمرونة ولديها القدرة على التكيف في الوقت المناسب وخاصة في التعاملات المالية والمصرفية، فمتغيرات السوق تحتاج إلى تصحيح مستمر وهذه الخوارزمية لها القدرة على التعامل مع هذه المهام والتغير في مؤشرات السوق(شيخة،٢٠٢٢).

تستخدم شجرة القرارات والشبكات العصبية في التنبؤ بالفشل والتعثر المالي، بالإضافة إلى تقدير خطر الائتمان وتقدير الربحية وتحليل النسب المالية، وتحليل مسببات التكاليف وكذلك تحليل إنتاجية القسم وتحليل الاحتيال الإداري(Kirkos ,et al., 2004; Sharma, et al., 2012) .

٢) الشبكات العصبية الاصطناعية:

إن فكرة عمل الشبكات العصبية الاصطناعية هو محاكاة البيانات للوصول إلى نموذج لهذه البيانات لغرض التحليل أو التصنيف أو التنبؤ، كما أنها تتميز بالمرونة العالية بالمقارنة

مع الأساليب الأخرى المستخدمة على نموذج البيانات وتخزين المعلومات وبثها في الشبكة العصبية الاصطناعية. وهي أكثر تكلفة من الناحية الحاسوبية من نظرائها في شجرة القرار.

ووفقا لدراسة (Digalaki,2019) فإن الشبكات العصبية الاصطناعية لديها إمكانات كبيرة لتوليد قيمة إضافية مقارنة بالمجالات التي تستخدم أدوات تكنولوجيا المعلومات الأخرى مثل الانحدار والتقدير والتجميع، ويقدرون أيضا أنه في ٦٩% من حالات الاستخدام الخاصة بهم، كما أنها توفر تحسينات تدريجية على التقنيات المستخدمة بالفعل، في حين أن ١٦٪ من الحالات تقدم الشبكات العصبية نتائج لا يمكن لأي تقنية تحليلية أخرى تقديمها.

ويعمل التصنيف هنا باستخدام بيانات منفصلة وغير مرتبة، ويساعد على تحديد تسميات الفئات الخاصة بأفراد المجموعة، ولكن نماذج التنبؤ تعمل بوظائف ذات قيمة مستمرة، وهذا يعني أنه يستخدم للتنبؤ بقيم البيانات الرقمية المفقودة أو غير المتوفرة من قيم سمات العينة.

يتم استخدام الشبكات العصبية وتحديدًا شبكة Multilayer Perceptron في اختيار المشروع الاستثماري الأمثل، مما يوفر أساس علمي مقبول من قبل كل الأطراف، فضلاً عن اختصار الجهد والوقت المبذولين في هذه القرارات (رمو،٢٠١٩). يتم استخدام الشبكات العصبية في التنبؤ بإجمالي الربح والتدفقات النقدية، والأداء المالي وتحليل النسب المالية من خلال الاعتماد على الشبكات العصبية (Amani, et al., 2017). وأيضاً في تحليل السلاسل الزمنية للتطبيقات المالية (Tarjoa, 2017) .

تقوم الشبكات العصبية على تصنيف المجموعات المتشابهة مع بعض، ثم يتم تصنيف الكائنات أو تجميعها بناء على مبدأ زيادة التشابه، تستخدم خوارزمية K-means بانتظام أبسط خوارزميات التجميع، حيث تنتج خوارزمية K-Means عدداً كبيراً من العناصر الخارجية فخريطة التنظيم الذاتي هي طريقة هامة قائمة على الشبكات العصبية تستخدم للتجميع وتم تطبيقها لحل المشاكل في القطاع المصرفي.

ساعدت الشبكات العصبية على التنبؤ بالتكاليف وتقديرها، كما تم استخدام شجرة القرارات لأغراض التصنيف، ويمكن استخدام الأسلوبين لتحديد أفضل تكلفة لكل بديل من خلال تحليل عميق للتكاليف واستنباط علاقات وأنماط جديدة.

٣) الغابات العشوائية :

هي تقنية التعلم الخاضع للإشراف، يمكن استخدامه لكل من مشاكل التصنيف والانحدار ، يعتمد على مفهوم التعلم الجماعي ، وهو عملية الجمع بين عدة مصنفات لحل مشكلة معقدة وتحسين أداء النموذج.

الغابات العشوائية هو مصنف يحتوي على عدد من أشجار القرار في مجموعات فرعية مختلفة من مجموعة البيانات المحددة، ويأخذ المتوسط لتحسين الدقة التنبؤية لمجموعة البيانات هذه. بدلاً من الاعتماد على شجرة قرار واحدة ، تأخذ الغابة العشوائية التنبؤ من كل شجرة وتعتمد على أغلبية الأصوات للتنبؤات ، وتتوقع الناتج النهائي، يؤدي العدد الأكبر من الأشجار في الغابة إلى دقة أعلى ويمنع مشكلة التجهيز الزائد.

وتتميز الغابات العشوائية بأنها تستغرق وقت أقل مقارنة بالأساليب الأخرى، والقدرة على التنبؤ بدقة أعلى، وبصفة خاصة بالنسبة لمجموعة البيانات الكبيرة ذات الأبعاد المتعددة، يمكنها المعالجة والحفاظ على الدقة عند وجود نسبة كبيرة من القيم المفقودة في البيانات.

٤) تحليل الانحدار اللوجستي:

يعتبر الانحدار اللوجستي الثنائي أحد أهم تقنيات تحليل الانحدار، وهو أسلوب التنبؤ الأكثر شيوعاً، باعتباره أكثر النماذج ملائمة لوصف البيانات التي تكون الاستجابة فيها ثنائية، ويكون فيها المتغيرات التوضيحية ذات مستويات متعددة. حيث أن الانحدار اللوجستي يستخدم عدة متغيرات متوقعة والتي يمكن أن تكون رقمية أو فئوية. (Baader & Krcmar, 2018)

يندرج الانحدار اللوجستي ضمن الخوارزميات الخاضعة للتصنيف، وهو يتميز بسمعة طيبة في القطاع المصرفي، بسبب قدرته البارزة في الكشف عن الاحتيالات في القوائم المالية، والكشف عن الغش فى بطاقات الائتمان، وحل المشاكل التي يسببها الغش فى تأمين السيارات والغش فى الشركات، بالإضافة إلى الكشف عن الغش فى المحاسبة المالية(الجبلي،٢٠٢).

كما يساعد على التعبير عن العلاقة بين المتغيرات كدالة خطية أو غير خطية، والتنبؤ بالعديد من المشاكل العالمية المتصلة بالبنوك، مثل توقعات أسعار الأسهم، أو تسجيل الائتمان، وأحجام المبيعات ومعدلات فشل المنتجات، لأنها تعتمد على التفاعلات المعقدة لمتغيرات التنبؤ المتعددة ، وأكثر ما يميز استخدام نماذج الانحدار سرعة استخراج النتائج وإمكانية الحصول على درجة الثقة عالية فى النموذج ، فضلا عن إمكانية التعامل مع عدد كبير من المتغيرات فى وقت واحد.

لذلك، قد تكون التقنيات الأكثر تعقيدًا (مثل الانحدار اللوجستي أو أشجار القرار أو الشبكات العصبية) ضرورية للتنبؤ بالقيم المستقبلية

وترى الباحثة أنه يمكن استخدام التنقيب في البيانات في المجال المصرفي من خلال تجميع وتصنيف العملاء، الذين لديهم نفس النوع من المعاملات أو الاستفسارات أو التوصيفات، أو الاشتراك في منتجات مماثلة أو لديهم نفس الاستعداد للمخاطر. فعلى سبيل المثال: يميل العملاء الذين يتقاضون أجوراً إلى الانضمام إلى خطط الاستثمار بمساهمات منتظمة، وتساعد المعرفة بهذه الحصص البنوك على تصميم منتجات لكل فئة من فئات العملاء، كما يمكنها البدء في حملات تسويقية هادفة وأكثر فعالية.

ومما سبق يتضح للباحثة أن التنقيب في البيانات DM له أثر كبير على إثراء مجالي المحاسبة والمراجعة بصفة عامة من خلال تحقيق أهدافه واستخدام تقنياته ، وله أثر كبير وفعال وضروري على القطاع المصرفي حيث يساعد في تكوين مجموعات وتصنيف وتقسيم العملاء والتنبؤ بالربحية، وتحليل واتخاذ قرار الائتمان، وتوقع الخطأ في الدفع والتسويق والمعاملات الاحتيالية، وترتيب الاستثمارات بصورة سريعة وفعالة، وتحسين محافظ الأسهم وإدارة النقد وعمليات التنبؤ، وتقييم مقدمي طلبات القروض عالية المخاطر، وعملاء بطاقات الائتمان الأكثر ربحية والبيع المتبادل، وذلك من خلال تحقيق هدفي التنبؤ التلقائي للاتجاهات والسلوكيات المستقبلية، والاكتشاف التلقائي لأنماط غير معروفة سابقاً.

كما تقوم تقنية التنقيب في البيانات DM باكتشاف المعاملات الشاذة، حيث تقارن المعاملات الجديدة بأنماط السلوك الشائعة لدى المستخدم، والمعاملة التي لا تتفق مع أحد هذه الأنماط ستعامل على أنها نشاط مريب وتطلق خطوات تحوطية وفقا لذلك.

كما يمكن استخدام تقنيات التنقيب في البياناتDM محللي المصارف من تحليل الاتجاهات السابقة وتحديد الطلب الحالي والتنبؤ بسلوك العملاء لمختلف المنتجات والخدمات من أجل اغتنام المزيد من الفرص التجارية وتوقع أنماط السلوك، تساعد تقنية جمع البيانات أيضاً على التعرف على العملاء المربحون من غير المربحين، بالإضافة إلى تقديم البنوك عرضاً جذاباً لعملائها عن طريق مطالبتهم بشراء منتجات أو خدمات إضافية.

٣/ الدراسة التطبيقية

تعتمد الباحثة لإجراء الجانب التطبيقي على دراسة حالة ، حيث تستطيع دراسة الحالة التركيز على الواقع العملي وتوفير مصادر متعددة للبيانات. كما أن النتائج التي توفرها دراسة الحالة تتسم بالعمق والجودة، وسوف تعتمد هذه الدراسة على تحقيق أهداف الوصف والتنبؤ لإعطاء إدارة البنك مؤشر مستقبلي لقدرة العميل على سداد القرض أم لا وفي الوقت المحدد لمساعدته في اتخاذ القرار، من خلال التركيز على العملاء الموثوق بهم واستبعاد غير الموثوق بهم، تجنباً لما يمكن أن يتحمله البنك من خسائر ناجمة عن عدم السداد.

ونظراً لما واجهته الباحثة من نقص شديد فى قواعد البيانات الواقعية التي يمكن استخدامها لإجراء الجانب التطبيقي، ونظرا لاستحالة الحصول على بيانات من البنوك المصرية عن قروض العملاء لسرية هذه البيانات لجأت الباحثة إلى الحصول على بيانات من موقع www.Kaggle.com وهي بيانات حقيقية متاحة للدراسة.

١/٣ التعريف بالبرنامج المستخدم Anaconda:

قامت الباحثة باستخدام برنامج اناكوندا Anaconda وهو برنامج مفتوح المصدر للغات البرمجة Python و R المستخدمة على نطاق واسع في الحوسبة العلمية (علوم البيانات، والتعلم الآلي Machine Learning، التعلم العميق Deep Learning ، والتحليلات التنبؤية، والبيانات الضخمة ، إلخ).

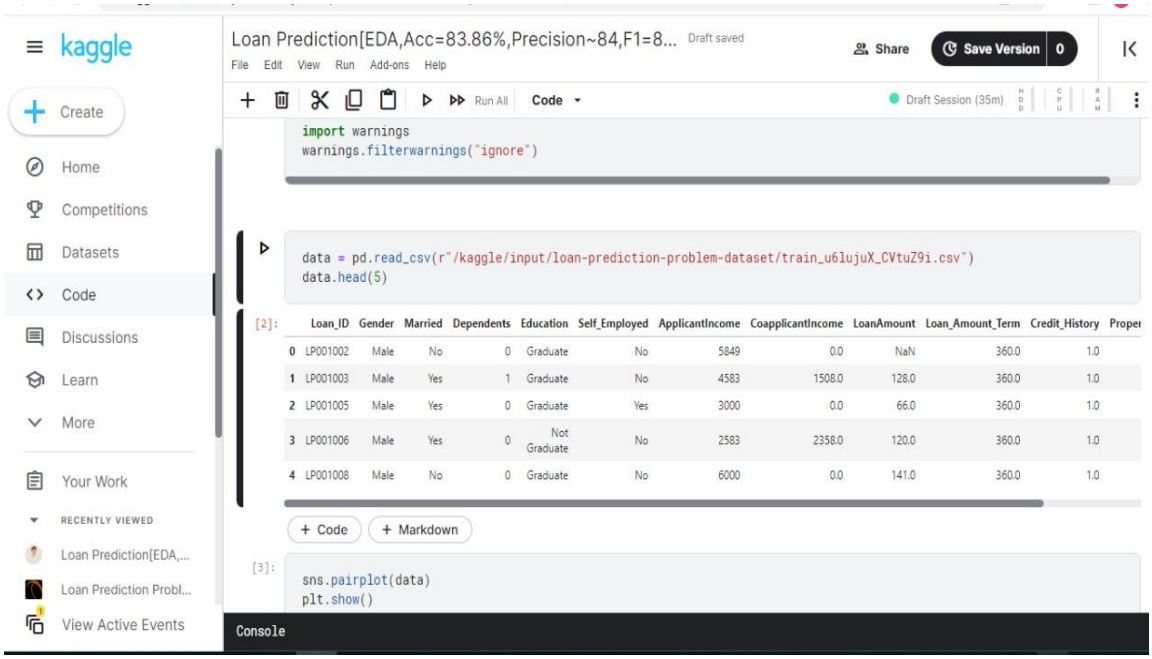
يمدنا البرنامج بمجموعة من خوارزميات تعلم الآلة Machine Learning والتي تستخدم لإجراء مهام تنقيب البيانات، ويعتمد البرنامج على توافر ملفين ملف اختبار file Test وملف البيانات التي تفحص Data file وكليهما بصيغة Excel sheet وذلك لاتخاذ القرارات المستقبلية استنادا على الواقع العملي لخبرات إدارة البنك في التعامل مع المقترضين السابقين وقدرتهم على السداد من عدمه.

وقد تم التركيز على استخدام لغة بايثون Python باعتبارها أحد لغات البرمجة المستخدمة لاستخدام أساليب وأدوات التنقيب في البيانات DM مثل : الانحدار اللوجستي Logistic Regression وشجرة القرارات Decision Tree و الغابات العشوائية Random Forest ...إلخ.

تقتصر الباحثة في هذه الدراسة عند استخدام الأساليب المختلفة للتنقيب في البيانات DM على اختبار المتغيرات التي تخص عملية قياس وتحليل والتنبؤ بمخاطر القروض والقدرة على السداد، حيث تم اختيار متغيرات البحث وفقا لبيانات العملاء التي تم الاعتماد عليها سابقاً من قِبل البنوك لاتخاذ قرارات إعطاء القروض.

ويتم العمل في البرنامج كما يلي :

I. إدخال متغيرات الدراسة إلى البرنامج كما يلي :



II. تهيئة البيانات والقيم المفقودة:

هنا يتم تنظيف (تصفية) البيانات (data cleaning) بمعنى التخلص من بعض البيانات التي تحتوي على عيوب أو بيانات قديمة لا تفيد في الوقت الحالي وغير ذلك. وقد قامت الباحثة بحذف البيانات المتضاربة وغير المهمة، واستبعاد بعض الصفات التي قد تكون ليست محل الدراسة.

القيم المفقودة (Missing value) بمعنى أن تكون البيانات كاملة ولا تحتوي على قيم مفقودة، وهنالك عدة طرق لإعداد وتصحيح البيانات، مثل الوسط الحسابي والانحراف المعياري.

III. اشتقاق ودمج البيانات(Data derivation& Merging):

في بعض الأحيان يستوجب اشتقاق بعض الأعمدة فتساعدنا في الحصول على معلومات مفيدة، ثم دمج البيانات بغرض تحقيق التكامل بين المصادر المتعددة للبيانات، حيث يتم دمج بعض الأعمدة للحصول على أفضل النتائج و اختصار البيانات. وقد قامت الباحثة في هذه المرحلة بدمج بيانات الفروع المختلفة من البنك فى قاعدة بيانات واحدة أو مخزن بيانات واحد لعمل مقارنة وجمع أكبر كمية من المعرفة للفروع مجتمعة.

IV. اختيار وتحويل البيانات (Data selection & transformation):

أي استرجاع كل البيانات التي لها علاقة بالدراسة من قاعدة البيانات. وفي هذه المرحلة تم تحديد البيانات المراد عمل التنقيب عليها. و تحويل البيانات إلى نماذج مخصصة للتنقيب .

V. تنقيب البيانات Data Mining:

أي استخدام أساليب وأدوات التنقيب في البيانات DM مثل : الانحدار اللوجستي Logistic Regression وشجرة القرارات Decision Tree و الغابات العشوائية Random Forest ...إلخ.

٢/٣ نتائج التحليل:

١/٢/٣ المؤشرات الاحصائي: (التحليل الوصفي لمتغيرات الدراسة)

يعتبر التحليل الاحصائي من أهم الأدوات التي يستخدمها الباحثين في تحليل البيانات من أجل الوصول إلى قرارات صحيحة عن طريق نتائج تحليل البيانات، لذلك يتم إجراء الأساليب الاحصائية من أجل التعرف على طبيعة البيانات والتعرف عليها . سيتم في هذا المبحث التحليل الاحصائي لبيانات الدراسة المتمثلة بالمتغيرات التالية (الجنس Gender، الحالة الاجتماعية Married، التعليمEducation، هل يعول Dependents، الحالة الوظيفية Self Employed، امتلاك بطاقة سابقاً Credit History، قيمة القرض Loan Amount، قسط القرض Loan Amount term، دخل العميل المشترك Co-applicant Income، دخل العميل الفردي Applicant Income، حالة قبول القرض Loan status)

جدول رقم(١)

يبين التكرارات والنسب المئوية لمتغيرات الدراسة

Variables		Frequency	Percent
Gender	Female	123	20%
	Male	491	80%
Married	Not-married	215	35%
	Married	399	65%
Education	Not-Graduate	134	21%
	Graduate	480	78%
Self-Employed	No	512	83.4%
	Yes	102	16.6%
Credit History	Not found	106	17.3%
	Found	508	82.7%
Loan Status	No	192	31.3%
	Yes	422	68.7%

من خلال جدول رقم (١) يتبين أنه بقياس نسب التكرارات والنسب المئوية للمتقدمين للحصول على قرض من البنك وفقاً لمتغيرات الدراسة:

نسبة الذكور إلى الإناث بلغت 80% : 20 %
نسبة المتزوجين إلى غير المتزوجين بلغت 65% : 35 %
نسبة المتعلمين إلى غير المتعلمين بلغت 78.2 % : 2.78%
نسبة الذين يشغلون وظيفة إلى الذين لا يشغلون وظيفة بلغت 16.6% : 83.4 %
نسبة الذين يمتلكون بطاقة بنكية إلى الذين لا يمتلكون بطاقة بنكية بلغت 82.7%: 17.3 %
نسبة الذين حصلوا على قرض إلى الذين لم يحصلون على قرض بلغت 68.7% : 31.3%

مما سبق يتضح أن أغلب الذين حصلوا على القرض من الذكور المتعلمين الذين لديهم بطاقة بنكية

كما قامت الباحثة بقياس المتوسط الحسابي والوسيط والانحراف المعياري لمتغيرات الدراسة القابلة للقياس كالتالي:

جدول (٢) قياس المتوسط الحسابي والوسيط والانحراف المعياري

Variables	Total	Min	Max	Mean	Std. Deviation
Loan Amount	614	9	700	146.5928	85.62432
Loan Amount term	614	12	480	41.1401	66.27544
Applicant Income	614	150	81000	5403.459	6109.0416
Co-applicant Income	614	0	41667	1621.245	2926.2483

من خلال جدول رقم(٢) السابق والذي يبين حساب المؤشرات الاحصائية، لاحظت الباحثة أن :

- ☒ أن قيمة متغير(Loan Amount) قد كانت أقل قيمة له هي (9) وأعلى قيمة له (700) في حين بلغت قيمة الوسط الحسابي له (146.5928) بانحراف معياري قدره (85.62432).
- ☒ أن قيمة متغير(Loan Amount term) قد كانت أقل قيمة له هي (12) وأعلى قيمة له (480) في حين بلغت قيمة الوسط الحسابي له (341.1401) بانحراف معياري قدره (66.27544).
- ☒ متغير(Applicant Income) قد كانت أقل قيمة له هي (150) وأعلى قيمة له (81000) في حين بلغت قيمة الوسط الحسابي له (5403.459) بانحراف معياري قدره (6109.04167).
- ☒ أن قيمة متغير(Co applicant Income) قد كانت أقل قيمة له هي (0) وأعلى قيمة له (41667) في حين بلغت قيمة الوسط الحسابي له (1621.245) بانحراف معياري قدره (2926.24837).

٢/٢/٣ قياس التداخل الخطي بين المتغيرات:

قامت الباحثة باستخدام مقياس Collinearity Diagnostics للتحقق من وجود أو عدم وجود مشكلة التداخل الخطي بين المتغيرات والتي تؤثر على اتخاذ القرار، حيث يقوم المقياس بحساب معامل Tolerance لكل متغير من المتغيرات المستقلة والتابعة مع إيجاد معامل تضخم التباين Variance Inflation Factor (VIF) لتلك المتغيرات، وتعتمد قاعدة اتخاذ القرار على

قيمة VIF حيث يعتبر النموذج خالي من مشكلة التداخل الخطى إذا كانت قيمة VIF أقل من 10% وقيمة معامل Tolerance أكبر من 0.1. كما أشارت بعض الدراسات إلى أن المتغيرات خالية من مشكلة التعدد الخطي إذا كانت قيمة اختبار معامل(VIF) أقل من 5%.

جدول رقم(٣)
دراسة التداخل الخطي بين المتغيرات باستخدام
معامل تضخم التباين (VIF) ومعامل Tolerance

Variables	Collinearity Statistics	
	Tolerance	VIF%
Gender	0.878	1.139
Married	0.866	1.154
Education	0.948	1.054
Self Employed	0.977	1.023
Applicant Income	0.631	1.584
Co-applicant Income	0.878	1.139
Loan Amount	0.610	1.638
Loan Amount Term	0.960	1.042
Credit History	0.993	1.007

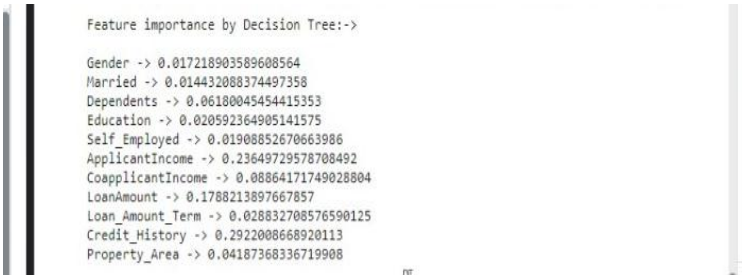
من خلال جدول رقم(٣) نلاحظ أن جميع المتغيرات المستقلة قد حصلت على معامل(VIF) لم يتجاوز 10%، كما أن معامل Tolerance لجميع متغيرات الدراسة أكبر من (0.1)، ويدل هذا على عدم وجود مشكلة التداخل الخطي بين المتغيرات المستقلة محل الدراسة.

٣/٣ تقييم أساليب التنقيب في البيانات:

هنالك عدة أساليب في تصنيف البيانات لا سيما التي تدخل في تنقيب البيانات وهي أسلوب شجرة القرارات، وأسلوب الغابات العشوائية، والانحدار اللوجستي .

١/٣/٣ شجرة القرار Microsoft Decision Tree :

يعتبر أسلوب شجرة القرارات أحد أساليب تنقيب البيانات، والذي بصورة أساسية لغرض التنبؤ، ويمكن أن يساعد أيضاً في توضيح كيفية توزيع البيانات (Seidman، 2001) أي أنه يسهل رسم استنتاجات عامة حول مجتمع البيانات.



جدول رقم(٤)

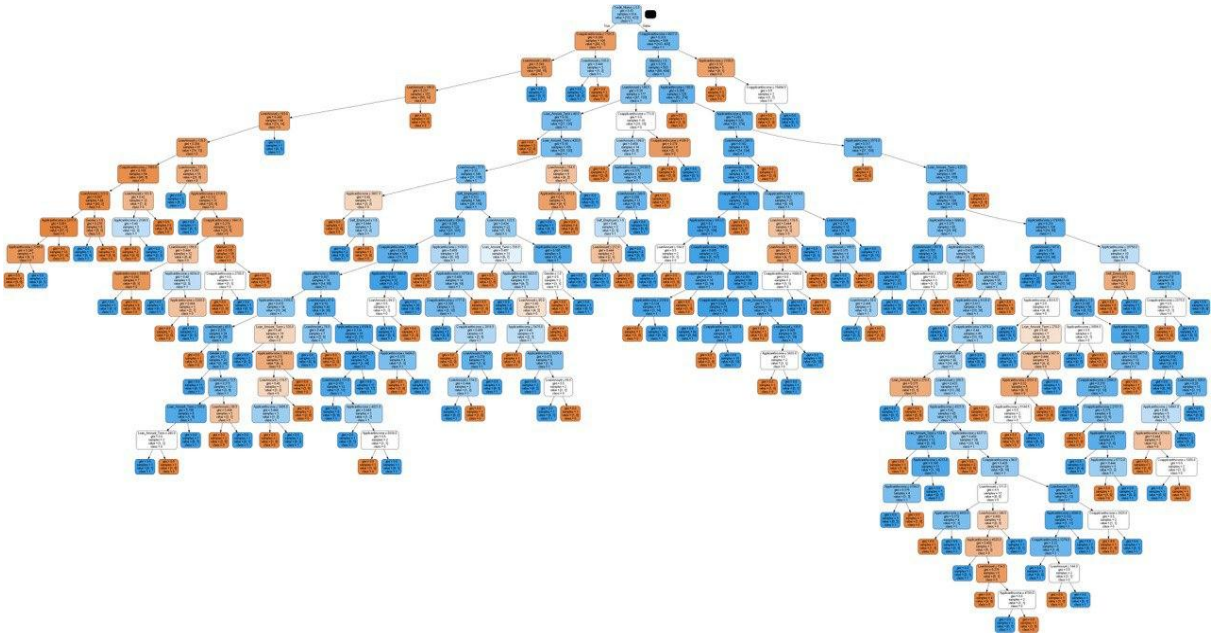
نتائج مخرجات شجرة القرارات

Variables	Importance	Accuracy
Gender	0.018754	81%
Married	0.026003	
Dependents	0.044890	
Education	0.023329	
Self Employed	0.012106	
Applicant Income	0.208975	
Co-applicant Income	0.114485	
Loan Amount	0.190074	
Loan Amount Term	0.043842	
Credit History	0.260012	
Property area	0.050562	

يظهر جدول رقم(٤) نتائج مخرجات شجرة القرارات والتي تظهر ترتيب أهمية المتغيرات المستقلة والتي تؤثر في عملية اتخاذ القرار لمنح القرض من قبل البنك، فتأتي أهمية المتغيرات كالتالي:

شكل رقم(٢)

يبين شجرة القرارات الخاصة بالتنبؤ بمنح القرض

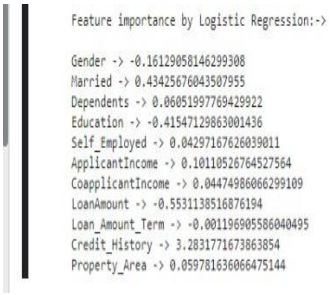


المرتبة الأولى: المتغير (Credit History) يمتلك أعلى أهمية إذ بلغ (0. 260012) المرتبة الثانية: المتغير (Applicant Income) إذ بلغت أهميته (0.208975) .
المرتبة الثالثة: المتغير(Loan Amount) إذ بلغت أهميته (0.190074)
المرتبة الرابعة: المتغير(Co-applicant Income) إذ بلغت أهميته(0.114485)

أما بقية المتغيرات فلم تكن ذات أهمية في عملية التنبؤ بمنح القرض أم لا لأن مقدار الأهمية Importance كان ضعيف جداً. في حين بلغت قيمة Accuracy من خلال استخدام شجرة القرارات للتنبؤ بـ(92%).

٢/٣/٣ الانحدار اللوجستي Logistic Regression:

تم استخدام أسلوب الانحدار اللوجستي في تحليل متغيرات الدراسة من أجل الحصول على أفضل النتائج، إذ يعتبر هذا الأسلوب هو الأساس في تصنيف البيانات.



```
Feature importance by Logistic Regression:->

Gender -> -0.16129058146299308
Married -> 0.43425676043507955
Dependents -> 0.06051997769429922
Education -> -0.41547129863001436
Self_Employed -> 0.04297167626039011
ApplicantIncome -> 0.10110526764527564
CoapplicantIncome -> 0.04474986066299109
LoanAmount -> -0.5531138516876194
Loan_Amount_Term -> -0.0011969055860404 95
Credit_History -> 3.2831771673863854
Property_Area -> 0.059781636066475144
```

جدول رقم(٥)

يبين نتائج مخرجات الانحدار اللوجستي

Variables	*Importance*	*Accuracy*
Gender	**0.016129**	**77%**
Married	**0.434256**	
Dependents	**0.060519**	
Education	**0.415471**	
Self Employed	**0.042971**	
Applicant Income	**0.201105**	
Co-applicant Income	**0.104285**	
Loan Amount	**0.190074**	
Loan Amount Term	**0.000119**	
Credit History	**3.283177**	
Property area	**0.059781**	

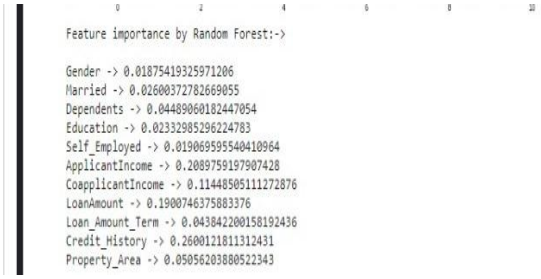
من خلال جدول رقم (٥) يبين نتائج مخرجات الانحدار اللوجستي والتي تظهر أهمية المتغيرات المستقلة والتي تؤثر في عملية اتخاذ القرار لمنح القرض من قبل البنك، إذ نلاحظ أن :

المرتبة الأولى: المتغير (Credit History) يمتلك أعلى أهمية إذ بلغ (3.283177)
المرتبة الثانية: المتغير (Applicant Income) إذ بلغت أهميته (0.201105)
المرتبة الثالثة: المتغير(Loan Amount) إذ بلغت أهميته (0.190074)
المرتبة الرابعة: المتغير(Co-applicant Income) إذ بلغت أهميته (0.104285)

أما بقية المتغيرات فلم تكن ذات أهمية في عملية التنبؤ بمنح القرض أم لا لأن مقدار الأهمية Importance كان ضعيف جداً. في حين بلغت قيمة Accuracy من خلال استخدام الانحدار اللوجستي للتنبؤ بـ(77%)

٣/٣/٣ الغابات العشوائية Random Forest:

تم استخدام الغابات العشوائية في تحليل متغيرات الدراسة من أجل الحصول على أفضل النتائج، إذ يعتبر هذا الأسلوب هو الأساس في تصنيف البيانات.



جدول رقم(٦)

نتائج مخرجات الغابات العشوائية

Variables	Importance	Accuracy
Gender	**0.017218**	**92%**
Married	**0.014432**	
Dependents	**0.061800**	
Education	**0.020592**	
Self Employed	**0.019088**	
Applicant Income	**0.236497**	
Co-applicant Income	**0.088641**	
Loan Amount	**0.178821**	
Loan Amount Term	**0.028832**	
Credit History	**0.292200**	
Property area	**0.041873**	

من خلال جدول رقم(٥) يبين نتائج مخرجات الغابات العشوائية Random Forest، والتي تظهر أهمية المتغيرات المستقلة والتي تؤثر في عملية اتخاذ القرار لمنح القرض من قبل البنك، اذ نلاحظ أن :

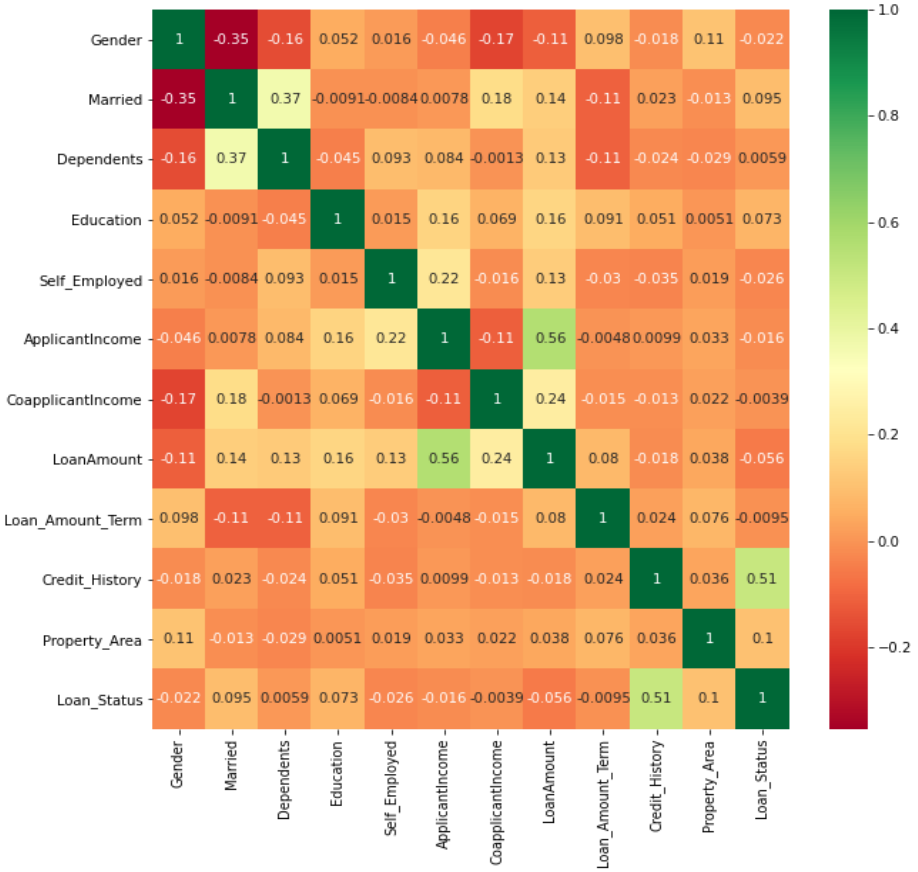
المرتبة الأولى: المتغير (Credit History) يمتلك أعلى أهمية إذ بلغ (3.283177)
المرتبة الثانية: المتغير (Applicant Income) إذ بلغت أهميته (0.201105)
المرتبة الثالثة: المتغير(Loan Amount) إذ بلغت أهميته (0.190074)
المرتبة الرابعة: المتغير(Co-applicant Income) إذ بلغت أهميته(0.104285)

أما بقية المتغيرات فلم تكن ذات أهمية في عملية التنبؤ بمنح القرض أم لا لأن مقدار الأهمية Importance كان ضعيف جداً .

في حين بلغت قيمة Accuracy من خلال استخدام الانحدار اللوجستي للتنبؤ بـ(77%)

يوضح الشكل رقم (٢) يوضح درجة الارتباط بين المتغيرات المختلفة والتي تم الاشارة اليها في السابق. يتضح أن تاريخ الائتمانCredit History هو الأعلى في الارتباط، دخل مقدم الطلب المشترك Applicant Income سلبي في الارتباط، ويمكن تفسيره، أن عدم وجود دخل للمقترض يزيد من المخاطر.

الشكل رقم (٢)
درجة الارتباط بين المتغيرات المختلفة



٤/٣ مقارنة بين دقة تطبيق أساليب التنقيب في البيانات DM الثلاثة:

وزيادة في قياس دقة النموذج التنبؤي لأهميته وأثره على الحد من المخاطر التي تواجهه البنوك عند منح العملاء القروض قامت الباحثة باختبار تحليل الحساسية من خلال تحليل منحنى ROC على أساليب التنقيب في البيانات الثلاثة التي تم اختبارها لتحديد أدقها نظراً لحساسية القرار الذي تتنبأ به.

تحليل منحنى Receiver Operating Characteristic(ROC): "هو مقياس لتقييم النواحي الإيجابية الخاطئة مقابل المبادلات السلبية الخاطئة، وأقل قيمة يأخذها صفر وأكبر قيمة واحد، ويكون التصنيف أفضل كلما اقتربت قيمة هذا المقياس من الواحد، وإذا انخفضت قيمة هذا المقياس عن 0.5، فإن ذلك سوف يؤثر بشكل سلبي على نموذج التصنيف" .

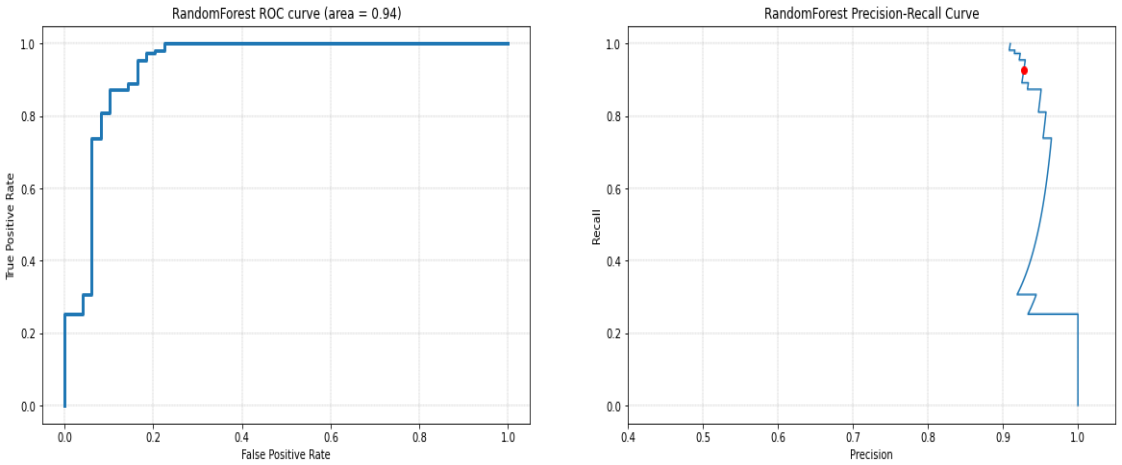
وقد قامت الباحثة بالاعتماد علي تحليل منحني الدقة (ROC) لقياس قدرة النماذج المقترحة علي التنبؤ الصحيح بحالة المتقدم بطلب القرض واتخاذ القرار السليم بالموافقة أو

الرفض. فعندما يكون الـ ROC بين 0.5 <ROC> 1، هناك احتمال كبير أن يكون المصنف قادرًا على التمييز بين قيم الفئة الموجبة وقيم الفئة السلبية. علي الجانب الآخر عندما تكون قيمة ROC = 0.5، فإن المصنف غير قادر على التمييز بين نقاط الفئة الإيجابية والسلبية. بمعنى إما أن المصنف يتنبأ بفئة عشوائية غير صحيحة أو فئة ثابتة لجميع نقاط البيانات. لذلك ، كلما ارتفعت قيمة ROC للمصنف، كانت قدرته أفضل على التمييز بين الفئات الإيجابية والسلبية.

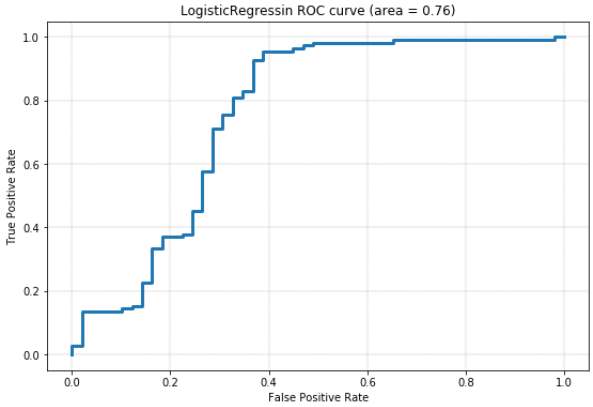
يتضح من الشكل رقم (3) أن منحنى الدقة Random Forest (ROC) يساوي 0.94 وهي نسبة جيدة جداً ومطمئنة للنموذج التنبؤي. كما أنا منحنى LR(ROC) يساوي 0.76 وهي مقبولة للنموذج التنبؤي، بينما منحنى الدقة Decision Tree(ROC) يساوي 0.82 معدل مقبول بنسبة كبيرة . وهكذا ترى الباحثة من خلال ما سبق قدرة نموذج Random Forest علي تحقيق أعلي معدل دقة مقارنة بنموذج Decision Tree ويأتي في المرتبة الثالثة من الدقة نموذج الانحدار اللوجستي LR .

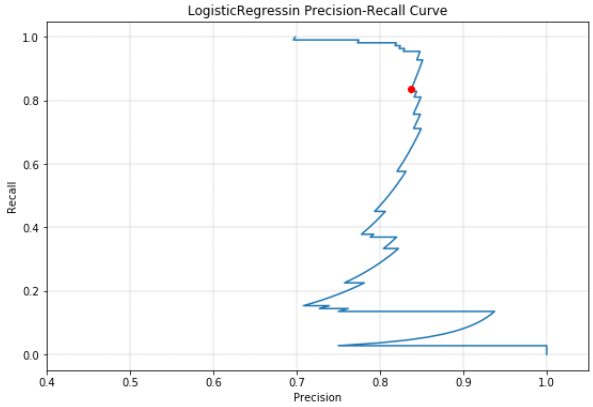
شكل رقم (٤)

Random Forest ROC (أ)

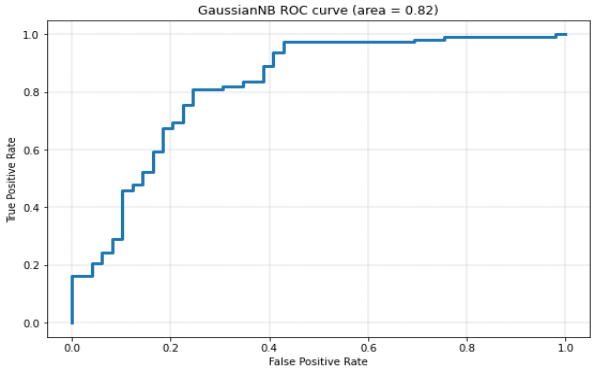


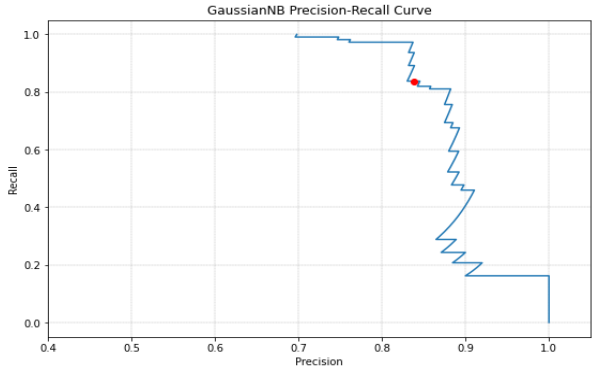
Logistic Regression ROC (ب)





Decision Tree ROC (ج)





كما يتضح من الجدول التالي، تتمتع الغابات العشوائية بأعلى درجات الدقة لذلك سيكون هذا هو النموذج الذي يجب استخدامه للتنبؤ بالحالة.

Model		Score positive	Score negative
Random Forest	Actual positive	181	0
	Actual negative	13	11
	Accuracy 0.94		
Logistic Regression	Actual positive	179	3
	Actual negative	13	10
	Accuracy 0.67		
Decision Tree	Actual positive	174	7
	Actual negative	9	15
	Accuracy 0.82		

٥/ الخلاصة والنتائج والتوصيات والأبحاث المستقبلية

٥/١الخلاصة

تقييم الائتمان: هو دراسة لتحديد جدوى طلب القرض الذي يتم تنفيذه، لتقييم ما إذا كان المقترض المحتمل لديه أنشطة تجارية ممكنة، قابلة للتسويق، مربحة، ومربحة يمكن دفعها على أساس الوقت. عادة ما يقوم موظف الحساب المصرفي الذي يمكن أن يكون جزءاً من لجنة التقييم– بتقييم الائتمان، ويتم إجراء هذا التقييم لتحليل جميع العوامل المشاركة في تطبيق الائتمان، مثل الأداء المالي للشركات والتصنيف الائتماني للمقرضين.

يساعد استخدام التنقيب في البيانات DM في القطاع المصرفي في تحقيق العديد من المزايا في مجال سرعة اتخاذ القرار، وتحقيق الميزة التنافسية.

كما تساعد تقنية التنقيب في البيانات DM من خلال استخدام تقنية الأنماط المتسلسلة على تحليل أنماط العملاء المختلفة، والمعاملات التي تؤدي إلى الاحتيال، مما يوفر المزيد من الوقت والجهد للكشف عن الغش والاحتيال. من المهم جداً في الكشف عن الاحتيال أن يتم العثور على المعاملات التي لا يقوم بها العميل، فنظام التنقيب في البيانات DM يقوم على التحقق من أي من هذه المعاملات لا يندرج ضمن فئة محددة أو لا يشكل معاملاً معيارياً مكرراً.

ومن ناحية أخرى، استخدام أساليب توقع القيمة_على سبيل المثال_ بدلاً من تصنيف تطبيقات القروض الجديدة، كمحاولة للتنبؤ بالمبالغ الافتراضية المتوقعة لتطبيقات القروض الجديدة، والقيم الرقمية المتوقعة وبالتالي فإنها تتطلب تقنيات نمذجة يمكنها أن تأخذ البيانات الرقمية كمتغيرات مستهدفة (أو متوقعة) من خلال استخدام الشبكات العصبية والانحدار اللوجستي. ويعد أكثر أساليب جمع البيانات شيوعاً المستخدمة في تحديد سمات العملاء هي:

١–التجميع (وصفي)

٢–التصنيف (تنبؤي) والتراجع (تنبؤي)

٣–اكتشاف قاعدة الاقتران (وصفي) واكتشاف النمط التسلسلي (تنبؤي)

غالباً ما تقدم المصارف خدمات استثمارية لعملائها، فالاستثمار هو: القيام باستثمار الأموال في الأصول أو أحد البنود للربح/الدخل، وهناك عدد هائل من الأدوات المالية في السوق.

فمن هنا ظهرت أهمية استخدام تقنية التنقيب في البياناتDM من خلال بناء نموذج تنبؤي ناجح، وتصوير التقرير في صورة معلومات مفيدة للمستخدم يمكن أن تسهل عليه جمع البيانات، باستخدام قواعد بيانات كبيرة.

٢/٥ النتائج :

١. إن مفهوم التنقيب في البيانات مفهوم معاصر يتعامل مع كمية كبيرة من البيانات من خلال توظيف برمجيات متخصصة تساعد في التنبؤ المستقبلي ومن ثم القدرة على معالجة هذه البيانات واسترجاعها بالسرعة والدقة المطلوبة.

٢. تتمثل أهم أهداف التنقيب في البياناتDM في الوصف والتنبؤ والتصور.

٣. توصلت الدراسة إلى أن التنقيب في البيانات يتم على متغيرات رقمية وغير رقمية، وتظهر في النموذج بالضبط كما تم صياغتها في قاعدة البيانات مثل مبلغ التمويل الذي يحمل تقديرات مثل كبير أو صغير.

٤. توصلت الدراسة إلى أن التنقيب في البيانات يساعد موظفي القطاع المصرفي على سرعة اتخاذ قرارات القروض، من خلال مراجعة سجلات المقرض، وتحليل قدرته على سداد الأقساط. بالإضافة إلى التنبؤ بالخدمات المستقبلية التي يحتاج إليها العملاء.

٥. توصلت الدراسة إلى أن التنقيب في البيانات في القطاع المصرفي يساعد على تحقيق الأمان وكشف الاحتيال، حيث تتم مراقبة البيانات الثانوية الكبيرة مثل سجلات المعاملات وتحليلها لتعزيز الأمن المصرفي، وتمييز السلوك غير العادي، والأنماط التي تشير إلى ذلك الاحتيال أو التصيد الاحتيالي أو غسل الأموال.

٦. توصلت الدراسة إلى أن التنقيب في البيانات القطاع المصرفي يساعد على تحليل بيانات بطاقة الائتمان الداخلية بحرية يمكن الوصول إليها من قبل البنوك تمكن من تسجيل الائتمان ومنح الائتمان التي تشكل جزءًا من الأكثر شعبية أدوات لإدارة المخاطر وتقييم الاستثمار

٧. توصلت الدراسة إلى أن التنقيب في البيانات مفيدا للغاية في القطاع المصرفي، من أجل تحسين استهداف العملاء الجدد واكتسابهم، واكتشاف الاحتيال في الوقت الفعلي، وتوفير

منتجات قائمة على القطاعات، وتحليل أنماط شراء العملاء عبر الوقت للحفاظ على العلاقات بشكل أفضل. ستحصل البنوك التي أدركت فائدة التنقيب عن البيانات والتي تعمل على بناء بيئة لاستخراج البيانات لعملية صنع القرار الخاصة بها على فائدة كبيرة وتستمد ميزة تنافسية كبيرة في المستقبل.

٨. توصلت الدراسة إلى أن التنقيب في البيانات يساعد في التنبؤ باكتشاف المخاطر التي تهدد البنك وذلك من خلال تحديد الخصائص التي تشير إلى وجود مخاطر تحد من منح العميل القرض.

٩. تتبنى القطاعات المصرفية تكنولوجيات التنقيب في البيانات في مختلف المجالات وخاصة في تقسيم العملاء والربحية، والتنبؤات بأسعار المنتجات الاستثمارية، وتجارة سوق المال ، والكشف الاحتيالي عن المعاملات، والتنبؤات بالمخاطر بمختلف أنواعها، والتنبؤ التقريبي بالتسعير، وسداد القروض.

١٠. توصلت الدراسة إلى أن أساليب التنقيب في البيانات المختلفة تعمل على تحسين دقة التنبؤ وسرعة منح القروض المصرفية، وفقا لتصنيفاتها القائمة على التصنيف والتجميع والتنبؤ والتصور.

١١. كما توصلت الدراسة تباين مستوى دقة التنبؤ بين أساليب التنقيب في البيانات DM، حيث أن نموذج الغابات العشوائية Random Forest هو الأعلى دقة في التنبؤ وفقا لمنحنى ROC والأكثر حساسية يليه أسلوب شجرة القرار Decision Tree ويأتي في المرتبة الثالثة أسلوب الانحدار اللوجستي LR .

٣/٥ التوصيات:

ومن خلال النتائج السابقة توصي الباحثة بما يلي:

١) ضرورة تدريب محللي النظم في القطاع المصرفي على التعامل مع الأدوات التكنولوجية الحديثة، التي تساعد على التعامل مع البيانات الضخمة والمتشعبة في البنوك ، للمساعدة في سرعة اتخاذ القرار، والذي يؤثر بالتالي على إدراك القطاع المصرفي مبكراً لوضعه المالي.

٢) مزيد من الدراسات حول استخدام أساليب أخرى من أساليب التنقيب في البيانات لمواجهة مختلف المشاكل المحاسبية والمخاطر التي يواجهه القطاع المصرفي.

٣) محاولة دمج أكثر من أسلوب من أساليب التنقيب في البيانات للاستفادة من مزايا كل أسلوب والتغلب على عيوبه في التعامل مع البيانات.

٤) محاولة التعمق في دراسة واستخدام أدوات الذكاء الاصطناعي (مثل تعلم الآلة machine learning، التعلم العميق deep learning) لتحليل وفهم الكميات الهائلة من البيانات التي تمتلكها البنوك والتي تتغير بشكل مستمر.

٥) استخدام التنقيب في البيانات للتعامل مع صفقات التجارة الالكترونية والتعامل بالعملات المشفرة.

٤/٥ الأبحاث المستقبلية:

١) استخدام التصور المرئي data visualization للتنقيب في البيانات المصرفية.

٢) استخدام التنقيب في البيانات في استخراج عن مؤشرات التداول في البنوك.

٣) وضع اطار مقترح لدمج أكثر من أسلوب من أساليب التنقيب في البيانات.

المراجـــع

أولاً: قائمة المراجع العربية:

أبو الخير، أسامة أحمد محمد، (٢٠١٩). 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*، كلية التجارة، جامعة كفر الشيخ، ع٧، ٣٠٥– ٣٤٧.
http://search.mandumah.com/Record/1000789

أبو خزانة، إيهاب محمد، (٢٠٠٧). نموذج مقترح لقياس مخاطر الائتمان المصرفي بهدف تطوير الافصاح المحاسبي وتكوين مخصص الائتمان بالبنوك التجارية. *إدارة الائتمان*، بنك الاسكندرية .

البدوي، سعد البدوي المبارك، (٢٠١٧). *استخدام تقنيات تنقيب البيانات لاستكشاف أنماط مؤثرا ت التحصيل الاكاديمي لطلاب المرحلة الثانوية* ، كلية علوم الحاسوب وتقانة المعلومات جامعة النيلين ، السودان

البشتاوي، سليمان حسين سليمان، البقمي، متعب عايش. (٢٠١٥). أثر تطبيق النظم الخبيرة في البنوك التجارية على اجراءات التدقيق الالكتروني من وجهة نظر المحاسبين القانونين الخارجيين. *المجلة الأردنية في إدارة الأعمال*، الجامعة الأردنية، عمادة البحث العلمي، مج ١١، ع١، –١٥١

الفارسي، أنوار فاضل، شحاتة، محمد موسى علي، و داود، ياسر إبراهيم محمد. (٢٠٢٢). دور أسلوب التنقيب في البيانات كأحد ابتكارات تكنولوجيا المعلومات في تحسين جودة التقارير المالية: دراسة ميدانية .*المجلة العلمية للدراسات والبحوث المالية والإدارية*، مج١٣، ع٢ ، ٦٢٨–٦٦٢.
http://search.mandumah.com/Record/1290999

العلى، عبد القادر، قنديل، عامر ابراهيم، الطمري، غسان، (٢٠٠٦)، "المدخل إلى دراسة المعرفة وإدارة السيرة"، عمان، ص: ١٧٥

الجابري، خالد محمد أحمد. (٢٠١٧). إدارة مخاطر الائتمان في البنوك التجارية اليمنية. *مجلة القلم*، ع٧، ٢٦٢-٢٨٢.
http://search.mandumah.com/Record/938087

الجبلى، وليد سمير عبدالعظيم، (٢٠٢٠). أثر استخدام الانحدار اللوجستي كأحد أساليب التنقيب في البياناتDM (Data Mining) في دعم الرأي المهني لمراجعي الحسابات- دراسة تطبيقية. *مجلة الإسكندرية للبحوث المحاسبية*، كلية التجارة، جامعة الاسكندرية، مج٤، ع ٢، ١-٦٠.
http://search.mandumah.com/Record/1070271

الحلو، جمال نعمان، معلمي، سيد مهدي. (٢٠٢٢). أهمية تنقيب البيانات في كشف الاحتيال في القوائم المالية. *مجلة الكلية الاسلامية الجامعة*، مج٢، ع ٦٧، ٣٠٢-٣٢٠.

تقارير البنك المركزي المصري. ٢٠١٨. ضوابط منح الائتمان: التصنيف الائتماني وأسس تقييم الجدارة الائتمانية للعملاء.
https://www.cbe.org.eg/ar/BankingSupervision/CCR/Pages/CentralCreditRegistry.aspx

تقارير البنك المركزي المصري. ٢٠٢٢. تجميع مخاطر الائتمان: التصنيف الائتماني وأسس تقييم الجدارة الائتمانية للعملاء.
https://www.cbe.org.eg/ar/BankingSupervision/CCR/Pages/CentralCreditRegistry.aspx

جوي، منى كاظم، و عليوي، عبير كاظم. (٢٠٢٢). تنقيب بيانات المصارف العراقية للفترة ٢٠١٥-٢٠١٨ باستخدام أسلوب التصوير المرئي .*مجلة الإدارة والاقتصاد*، ع١٣٢، ١١١-١٤٤.
http://search.mandumah.com/Record/1263492

حسين، محمود محمد عبدالرحيم. (٢٠٢١). أثر الإفصاح عن مخاطر الائتمان المصرفي على تحسين الأداء المالي للبنوك التجارية في بيئة الأعمال المصرية: دراسة تطبيقية، *المجلة العلمية للبحوث التجارية*، س٨، ع٤ ، ١٢٣-١٥٥.
http://search.mandumah.com/Record/1231261

حسين، محمود محمد عبدالرحيم. (٢٠٢١). أثر الإفصاح عن مخاطر الائتمان المصرفي على تحسين الأداء المالي للبنوك التجارية في بيئة الأعمال المصرية: دراسة تطبيقية. *المجلة العلمية للبحوث التجارية*، س٨، ع٤، ١٢٣-٢٠٥.
http://search.mandumah.com/Record/1231261

رمو، وحيد محمود، (٢٠١٩) .التنقيب المحاسبي عن البيانات باستخدام الشبكات العصبية حالة دراسية. *مجلة العلوم الاقتصادية والادارية*. كلية الادارة والاقتصاد ،جامعة بغداد، مج ٢٥، ع ١١.

شاهين، عبدالحميد أحمد ، و البغدادي، رجب محمد عمران أحمد. (٢٠١٨). القياس المحاسبي لمخاطر الائتمان في ضوء معايير الرقابة المصرفية لبازل ٩ والمعيار IFRS 9:دراسة ميدانية بالبنوك التجارية المصرية .*المجلة العلمية للدراسات والبحوث المالية والإدارية*، مج٤، ع٢،١٠٩-١٥٦.
http://search.mandumah.com/Record/1109403

شيخة، عبادة. (٢٠٢٢). استخدام تقنيات التنقيب في البيانات بهدف الزيادة من الربحية،
http://eresearch.hiba.edu.sy/sites/default/files/2022-

صالح، سمير أبو الفتوح، (٢٠١٦). تحسين جودة التقارير المالية باستخدام التنقيب في البيانات. *المجلة المصرية للدراسات التجارية*، كلية التجارة، جامعة المنصورة، مج٤٠، ع ٢، ٥٩٧ – ٥٩٩.
http://search.mandumah.com/Record/775682

عبدالغفار، نورهان السيد محمد، (٢٠٢٠). نتائج استخدام أسلوب التنقيب في البيانات لدعم المحتوى المعلوماتي للقوائم المالية المستقبلية وأثر ذلك على تعزيز كفاءة القرارات الاستثمارية في السوق المالي المصري. *مجلة البحوث المالية والتجارية*، كلية التجارة، جامعة بورسعيد، ع ١، ٣٧٦-٤٦٥.
http://search.mandumah.com/Record/1090245

علي، هدى عبدالرحيم حسين. (٢٠١٨). استخدام تقنية التنقيب عن البيانات لتحليل المؤشرات المالية لعينة من المصارف الأهلية العراقية باعتماد خوارزمية .CART ، *المجلة العراقية لتكنولوجيا المعلومات*، مج٩، ع٢ ، ٣٢-٥٤.
http://search.mandumah.com/Record/1028955

عوض، عمرو إبراهيم، (٢٠٢٠). إطار مقترح لتحقيق التكامل بين أسلوب تحليل النظم وتقنيات التنقيب في البيانات بهدف دعم المراجع الخارجي لاكتشاف الأخطاء الجوهرية بالقوائم المالية والتقرير عن استمرارية المنشأة، *المجلة العلمية للدراسات والبحوث المالية* والتجارية، كلية التجارة، جامعة دمياط، مج١، ع١، ٥٣-١٤٥.
http://publication.du.edu.eg/journal/ojs302design/index.php/comm/article/view/2402

قاسم، زينب عبدالحفيظ أحمد. (٢٠١٨). إطار مقترح للإفصاح عن المخاطر الائتمانية وانعكاسات ذلك على جودة التقارير المالية للبنوك: دراسة تطبيقية، *مجلة الدراسات الاقتصادية المعمقة*، ع٧ ، ٢٨- ٥١.
http://search.mandumah.com/Record/1142857

مبارز، أسامة محمد مهدى. (٢٠١٦). تكنولوجيا المعلومات ودورها في رفع كفاءة الأداء المصرفي باستخدام بطاقة الأداء المتوازن: دراسة ميدانية على البنوك التجارية المصرية، *المجلة العلمية للدراسات التجارية والبيئية*، كلية التجارة، جامعة قناة السويس، مج ٧، ملحق ع٣، ٧٣٥-٧٥٤.
http://search.mandumah.com/Record/862948

نجيب، دحدوح، و علي، دبي. (٢٠٢١). أثر تدابير إدارة المخاطر الائتمانية على الأداء المالي في البنوك التجارية الجزائرية: دراسة قياسية باستخدام نماذج البيانات الزمنية المقطعية للفترة ٢٠٠٩-٢٠١٨، *مجلة البحوث في العلوم المالية والمحاسبية*، مج٦، ع١ ، ٣٠٢-٣٢٠.
http://search.mandumah.com/Record/1200923

ثانياً: قائمة المراجع الأجنبية:

Ahmad, A. E. M. K., & Al-Zu'bi, H. A. (2011). E-banking functionality and outcomes of customer satisfaction: an empirical investigation. ***International journal of marketing studies, 3*****(1), 50-65.**

Al-Azmi, A. A. R. (2013). ***Data, text and web mining for business intelligence: a survey.arXiv***

Albashrawi, M., (2016), Detecting financial fraud using data mining techniques: A Decade Review from 2004 to 2015, *Journal of Data Science*, 14(3), Pp: 553-569

Al-Hashedi, K. G., & Magalingam, P. (2021). Financial fraud detection aPplying data mining techniques: A comprehensive review from 2009 to 2019. *Computer Science Review*, 40, 100402. https://doi.org/10.1016/j.cosrev.2021.100402

Amani, F. A., & Fadlalla, A. M. (2017). Data mining aPplications in accounting: A review of the literature and organizing framework. *International Journal of Accounting Information Systems*, *24*, 32-58.

Aziz, A., Saha, S., & Arifuzzaman, M. (2021). Analyzing Banking Data Using Business Intelligence: A Data Mining APproach. In: Uddin, M.S., Bansal, J.C. (eds) Proceedings of International Joint Conference on Advances in Computational Intelligence. Algorithms for Intelligent Systems. Springer, Singapore. https://doi.org/10.1007/978-981-16-0586-4_20

Baader G. & Krcmar H., (2018), Reducing false positives in fraud detection: combining the red flag aPproach with process mining, *International Journal of Accounting Information Systems*, Vol. 31, Dec., PP: 1-16

Barman, S., Pal, U., Sarfaraj, M. A., Biswas, B., Mahata, A., & Mandal, P., (2016), A complete literature review on financial fraud detection aPplying data mining techniques, International Journal of Trust Management in Computing and Communications, Vol.3, No.4, PP: 336-359.

Bhambri, V., (2011), APplication of data mining in banking sector, *IJCSt*, Vol. 2, Issue 2, Pp:199-202.

Bharati, M., & Ramageri, M., (2010), "Data mining techniques and aPplications", *Indian Journal of Computer Science and Engineering*, Dec, 1, 4 , PP:301-305

Bou-Hamad I. & I., Jamali, (2020), Forecasting financial time-series using data mining models: A simulation study, *Research in International Business and Finance* ,Vol. 51, Jan,PP:1-14

Carneiro, N., Figueira, G. & Costa, M., (2017), A data mining based system for credit-card fraud detection in e-tail, *Decision SuPport Systems*, Vol. 95, March, Pp: 91-101

Chahadah, A. A., Refae, G. A. E., & Qasim, A., (2018), The use of data mining techniques in accounting and finance as a corporate strategic tool: an empirical investigation on banks operating in emerging economies, *International Journal of Economics and Business Research*, 15, 4, Pp:442-452.

Chou, J.S., Tai, Y., &Chang, L.J., (2010), *Predicting the development cost of TFT-LCD manufacturing equipment with artificial intelligence models,* Int. J. Prod. Econ. (128), (1), Pp:339-350.

Chui S, & Ding, N., (2018) Construction of a bank customer data warehouse and an aPplication of data mining. *In: Proceedings of 10th international conferences on machine learning and computing,* Pp 161–166

Colye, Brian, (2000), *Measuring credit risk*, CIB publishing United Kingdom, 2000, p. 1.

Demirel, S., (2009), The use of data Mining Techniques in Detecting Fraudulent Financial Statements: An APplication on Manufacturing Firms, *The Journal of Faculty of Economics and Administrative Sciences*, Vol.14, No.2, Pp: 157-170.

Digalaki E, (2019), "AI in Banking Business Insider", https://www.businessinsider.fr/us/the-ai-inbanking-report-2019-6

Du, J., P., (2010), Predicting bankruptcy using neural networks and other classification methods: the influence of variable selection techniques on model accuracy, *Neuron computing* 73 (10), 2047–2060.

Ekinci, R., & Poyraz, G. (2019). The effect of credit risk on financial performance of deposit banks in Turkey. *Procedia Computer Science*, *158*, 979-987.

Embark ،A. S. ،Haggag ،R. Y. ،& Saleh ،S.A F.F (2022). Banking Risk Prediction Framework Using Machine Learning Techniques. *Journal of Theoretical and APplied Information Technology*, *100*(20).

Ernawati, E., Baharin, S. S. K., & Kasmin, F. (2021, April). A review of data mining methods in RFM-based customer segmentation. In *Journal of Physics: Conference Series* (Vol. 1869, No. 1, p. 012085). IOP Publishing.

Gray, G.L., Debreceny, R.S., (2014), A taxonomy to guide research on the aPplication of data mining to fraud detection in financial statement audits, Int. J. Account. Inf. Syst. 15 (4), Pp: 357–380

Gu, C. (2022). APplication of data mining technology in financial intervention based on data fusion information entropy. *Journal of Sensors*, 2022.

Hamid, A. J., & Ahmed, T. M. (2016). Developing prediction model of loan risk in banks using data mining. *Machine Learning and APplications: An International Journal (MLAIJ)*, *3*(1), 1-9.

Hassani,H., Huang X. , Silva E.,(2018), Digitalisation and Big Data Mining in Banking ,Big Data and Cognitive Computing, 2, (18) www.mdpi.com/jou

Hassoon, S., & Nabi, N. A. (2022). The effect of credit risk on financial performance in Iraqi private commercial banks: Bank of Baghdad as a model. *World Economics and Finance Bulletin*, *12*, 101-116.

Isamil E.A., Amer W.M., (2020), The Impact of Digital Financial Technology on Financial Inclusion in Banking Sector: Evidence from Africa, *The challenges and prospects of the accounting and auditing profession in the twenty-first century, The Fourth scientific conference of the Department of Accounting and Auditing*, Faculty of Commerce, Alexandria University, 24-25 Dec.

Jabari, A. & Suyunus, M., (2018), Fraud Detection: Using Of Data Mining In Accounting Information System, *In Proceedings*

of the Journal of Contemporary Accounting and Economics Symposium 2018 on Special Session for Indonesian Study - Contemporary Accounting Studies in Indonesia, **Pp: 23-30**

Kirkos, K.,Spathis c.,Manopoulos C.,(2007), Data Mining techniques for the detection of fraudulent financial statements, ***Expert Systems with APplications*** **.Vol. 32, Issue 4, May, Pp: 995-1003.**

larose, D.T., (2005), ***Discovering Knowledge in Data: An Introduction to Data Mining,*** **John wiley & Sons, Inc., Hobken, New Jersey, Pp: 11-17**

Mandala, I. G., Nawangpalupi, C. B.,& Praktikto F. R., (2012), Assessing Credit Risk: An APplication of Data Mining in a Rural Bank, ***Procedia Economics and Finance,*** **Vol. 4, Pp: 406-412.**

Miyan, M., (2017), APplications of Data Mining in Banking Sector, ***International Journal of Advanced Research in Computer Science,*** **8, Pp:108-114.**

Mogos, G. Jamail, N. S. M. (2021). Study on security risks of e-banking system. ***Indonesian Journal of Electrical Engineering and Computer Science, 21*****(2), 1065-1072.**

Ngai, E. W., Hu, Y., Wong, Y. H., Chen, Y., & Sun, X., (2011), The aPplication of data mining techniques in financial fraud detection: A classification framework and an academic review of literature, ***Decision suPport systems,*** **Vol.50, Issue 3, Pp:559-569.**

Pandimurugan, V., Usha, D., Guptha, M. N., & Hema, M. S. (2022). Random forest tree classification algorithm for predicating loan. ***Materials Today: Proceedings, 57,*** **2216-2222.**

Peng, H., Lin, Y., & Wu, M. (2022). Bank Financial Risk Prediction Model Based on Big Data. ***Scientific Programming*****.**

Perera, R. A. A. S., & Thrikawala, S. S. (2010). An empirical study of the relationship of accounting information to the investor? S decisions

Pulakkazhy, S. & Balan, R. V., (2013), Data mining in banking and its aPplications- A review, *Journal of Computer Science*, 9(10), Pp:1252-1259. https://doi.org/10.3844/jcssp.2013.1252.1259

Ramachandran, P., Girija, N., & Bhuvaneswari, T. (2011). Classifying Blood Donors Using Data Mining Techniques. *International Journal of Computer Science Engineering & Technology*, *1*(1).

Ravisankar, P., Ravi, V., Rao, G. R., & Bose, I. (2011). Detection of financial statement fraud and feature selection using data mining techniques. *Decision suPport systems*, 50(2), 491-500.

Riasi, A., & Wang, D. (2016). Comparing the performance of different data mining techniques in evaluating loan aPplications. *International Business Research*, *9*(7), 164-187.

Safari, S., & Soleimani, R. (2020). E-banking Operational Risk Management Analysis by Self-Assessment APproach (A Case Study: a State Bank). *Commercial Strategies*, *9*(47), 309-320.

Samuel, O. L. (2015). The effect of credit risk on the performance of commercial banks in Nigeria. *African Journal of Accounting, Auditing and Finance*, *4*(1), 29-52.

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), Pp. 37-47.

Tarjoa, Herawati N., (2017), APplication of Beneish M-Score Models and Data Mining to Detect Financial Fraud, *Procedia - Social and Behavioral Sciences,* 25 Nov., Vol. 211, Pp: 924 – 930.

Tiwari, M., (2010), 'Data mining: A competitive tool in retail industries'', Global J. Enterprise Inform. Syst.,

Xiao, M., Xiaoli, H., Gaojin, L., (2010), Research on aPplication of data mining technology in financial decision suPport system,

Proceeding of the International IEEE Conference on Information Management, **Innovation Management and Industrial Engineering (ICIII), Kunming, China Nov., Pp: 381–384.**

Yang, J.G.S., (2006), Data mining techniques for auditing attest function and fraud detection, ***Journal of Forensic Investigative Accounting,*** **1, (1), Pp:4–10**

Yigitbasioglu, O.M., Velcu, O., (2012), ***A review of dashboards in performance management: implications for design and research,*** **Int. J. Account. Inf. Syst. , 13, (1), Pp:41–59.**

Zarei, S. (2011). Risk management of internet banking. In ***10th WSEAS International conference on Artificial Intelligence, Knowledge Engineering and Data Bases,*** **Cambridge, UK, Pp. 134–139.**