



الجامعة الافتراضية السورية
SYRIAN VIRTUAL UNIVERSITY

الجمهورية العربية السورية
وزارة التعليم العالي
الجامعة الافتراضية السورية

دور التنقيب في البيانات في اتخاذ قرارات المنح الائتماني

دراسة حالة بنك الشام

The role of data mining on granting credit decisions

Case study Cham Bank

بحث مقدم لنيل درجة الماجستير في ادارة الأعمال التخصصي

إعداد الطالبة

ريم فريد محمد شيخي / reem_122557

إشراف

د. حيدر عبد الله

الإهداء

الحمد لله الذي أكرمني بفضله وأنار طريقي ودربي، الحمد لله الذي أعننا ووفقنا في إنجاز هذا العمل.

إلى سندي في الحياة..... والأقرب الى روجي أبي العزيز

إلى داعمتي والمؤمنة الأولى بقدراتي، الى مصدر الحب والحنان..... أمي الغالية

إلى من وقف بجانبني وتحمل مشاق الدراسة والتعب ومصاعب الحياة.....زوجي الغالي

إلى من استمد منهم قوتي وإصراري.....اخوتي

إلى الأخت والصديقة التي سارت معي خطوة بخطوة في مسيرتي العلمية صديقتي شهد

إلى من كانوا سبباً في دعم مسيرتي التعليمية والمهنية.....زملائي ببنك الشام

إلى جميع من تلقَّيتُ منهم النصح والدعم

أهديكم خلاصة جهدي العلمي

شكر وتقدير

أتقدم بجزيل الشكر والتقدير لكل من وقف معي وساندني خلال فترة دراستي في الجامعة الافتراضية السورية، وأخص بالذكر أستاذي العزيز، المشرف على الرسالة الدكتور حيدر عبد الله لكرم عطائه ومساندته المستمرة من خلال توجيهاته وملاحظاته التي ساهمت بشكل كبير في إنجاز هذا العمل.

كما أتقدم بالشكر لجميع أعضاء الكادر التدريسي في الجامعة الافتراضية السورية لما قدموه فترة دراستي من تسهيلات ودعم.

وأتقدم بوافر التقدير وعظيم الامتنان للجنة المناقشة الأفاضل الذين أثروا هذا البحث من خلال ملاحظاتهم وإرشاداتهم القيمة.

كما أتقدم بالشكر لمن لم يتسنى لي ذكرهم وكان لهم دور كبير في المساعدة في إنجاز هذا البحث حتى أبصر النور.

شكراً لكم جميعاً.

قائمة المحتويات

| | |
|----|---|
| 1 | الفصل الأول: الإطار المنهجي للدراسة |
| 1 | مقدمة: |
| 2 | أولاً-مشكلة البحث: |
| 3 | ثانياً-أهداف البحث: |
| 3 | ثالثاً-أهمية البحث: |
| 3 | 1-أهمية علمية: |
| 3 | 2-أهمية تطبيقية: |
| 4 | رابعاً-منهجية البحث: |
| 4 | خامساً-حدود البحث: |
| 4 | سادساً-الدراسات السابقة: |
| 13 | الفصل الثاني: الإطار النظري للبحث |
| 13 | أولاً: الانتمان المصرفي: |
| 13 | 1-مفهوم الانتمان المصرفي: |
| 14 | 2-أهمية الانتمان المصرفي: |
| 15 | 3-أشكال الانتمان المصرفي: |
| 17 | 4-العوامل المؤثرة في قرار منح الانتمان: |
| 22 | 5-مخاطر الانتمان المصرفي: |
| 28 | ثانياً: التنقيب في البيانات: |
| 28 | 1-مفهوم وأهمية التنقيب في البيانات: |
| 29 | 2-أهداف التنقيب في البيانات: |
| 30 | 3-مراحل عملية التنقيب في البيانات: |
| 32 | 4-أنواع التنقيب في البيانات: |
| 33 | 5-أدوات التنقيب في البيانات: |
| 37 | 6-تطبيقات التنقيب في البيانات: |
| 40 | 7-استراتيجيات نجاح التنقيب في البيانات: |
| 40 | 8-صعوبات التنقيب في البيانات: |
| 41 | ثالثاً: دور التنقيب في البيانات في اتخاذ قرارات منح الانتمان: |
| 47 | الفصل الثالث: الإطار العملي للبحث |
| 47 | أولاً: لمحة عن بنك الشام: |
| 48 | ثانياً: الأدوات والبرامج المستخدمة: |
| 49 | ثالثاً: التحليل والخوارزميات المستخدمة: |
| 49 | 1- تحليل المشكلة: |
| 49 | 2- جمع البيانات: |
| 49 | 3- عينة البحث: |
| 50 | 4-وصف وتحليل البيانات: |
| 50 | 5-تهيئة البيانات: |
| 51 | 1.5.تنظيف البيانات (DATA CLEANING): |

| | |
|----|--|
| 51 | 2.5. تحويل البيانات (DATA TRANSFORMATION): |
| 52 | 3.5. تمثيل قيمة السمة واشتقاقها (ATTRIBUTE VALUE REPRESENTATION AND DERIVATION): |
| 53 | 4.5. اختيار السمات (ATTRIBUTE SELECTION): |
| 54 | 6- الخوارزميات المستخدمة: |
| 55 | رابعاً: مرحلة التنفيذ: |
| 55 | 1- مرحلة بناء قاعدة البيانات: |
| 58 | 2- تهيئة بيئة العمل لاختيار الخوارزميات: |
| 61 | 3- مرحلة تطبيق خوارزمية شجرة القرار: |
| 75 | 4- مرحلة تطبيق خوارزمية بيبز المبسط (NAÏVE BAYES): |
| 86 | الفصل الرابع: النتائج والتوصيات |
| 86 | أولاً- النتائج: |
| 86 | ثانياً- التوصيات: |

قائمة الجداول

| رقم الصفحة | الموضوع | رقم الجدول |
|------------|---|------------|
| 52 | تمثيل سمات خصائص التمويل | 1 |
| 52 | تمثيل سمات خصائص الضمانات | 2 |
| 53 | تمثيل سمات خصائص العميل | 3 |
| 54 | تمثيل اختيار السمات | 4 |
| 62 | يوضح مصفوفة التعارض | 5 |
| 69 | مصفوفة التعارض لشجرة القرار عن بيانات 2018 | 6 |
| 70 | نتائج مقارنة العملاء الجدد 2019 مع شجرة القرار | 7 |
| 72 | مصفوفة التعارض لشجرة القرار عن بيانات 2018 و2019 | 8 |
| 73 | نتائج مقارنة العملاء الجدد 2020 مع شجرة القرار | 9 |
| 75 | مصفوفة التعارض لشجرة القرار عن بيانات 2018 و 2019 و2020 | 10 |
| 78 | جدول تفصيلي لملف السمات بالنسب عن بيانات 2018 | 11 |
| 79 | مصفوفة التعارض لخوارزمية بيبز المبسط عن بيانات 2018 | 12 |
| 79 | مقارنة العملاء الجدد 2019 مع نتائج خوارزمية بيبز المبسط | 13 |
| 80 | نتائج مقارنة العملاء الجدد 2019 مع بيبز المبسط | 14 |
| 81 | جدول تفصيلي لملف السمات بالنسب عن بيانات 2018 و2019 | 15 |
| 82 | مصفوفة التعارض لخوارزمية بيبز المبسط عن بيانات 2018 و2019 | 16 |
| 83 | مقارنة العملاء الجدد 2020 مع نتائج خوارزمية بيبز المبسط | 17 |
| 83 | نتائج مقارنة العملاء الجدد 2020 مع بيبز المبسط | 18 |
| 84 | جدول تفصيلي لملف السمات بالنسب عن بيانات 2018 و2019 و2020 | 19 |
| 85 | مصفوفة التعارض لخوارزمية بيبز المبسط عن بيانات 2018 و2019 و2020 | 20 |

قائمة الأشكال

| رقم الصفحة | الموضوع | رقم الشكل |
|------------|--|-----------|
| 32 | مراحل اكتشاف المعرفة | 1 |
| 32 | أنواع تنقيب البيانات | 2 |
| 55 | واجهة البرنامج المستخدم SQL SERVER | 3 |
| 56 | نشاء ملف جديد لقاعدة البيانات | 4 |
| 56 | واجهة خيارات انشاء ملف لقواعد البيانات | 5 |
| 57 | انشاء جدول ضمن قاعدة البيانات | 6 |
| 57 | تعريف وتحديد خصائص الحقول ضمن الجدول | 7 |
| 57 | ادخال البيانات الى الجدول | 8 |
| 58 | انشاء استعلام للبيانات | 9 |
| 59 | واجهة برنامج Visual Studio 2019 | 10 |
| 59 | واجهة اختيار نموذج التنقيب في البيانات | 11 |
| 59 | واجهة تسمية المشروع | 12 |
| 60 | واجهة Source New Data | 13 |
| 60 | آلية الوصل مع قاعدة البيانات | 14 |
| 60 | واجهة views Data Source | 15 |
| 61 | اختيار الجدول من قاعدة البيانات | 16 |
| 63 | بناء نموذج التنقيب | 17 |
| 63 | اختيار خوارزمية شجرة القرار | 18 |
| 64 | تحديد الواصفات | 19 |
| 64 | تحديد نسبة البيانات المراد اختبارها | 20 |
| 65 | ضبط متغيرات خوارزمية Microsoft Decision Trees | 21 |
| 66 | تأكيد بناء النموذج | 22 |
| 66 | اكمال بناء النموذج | 23 |
| 67 | تنفيذ النموذج | 24 |
| 68 | شجرة القرار عن بيانات 2018 | 25 |
| 70 | مقارنة العملاء الجدد 2019 مع قواعد شجرة القرار | 26 |
| 71 | شجرة القرار عن بيانات 2018 و2019 | 27 |
| 72 | مقارنة العملاء الجدد 2020 مع قواعد شجرة القرار | 28 |
| 74 | شجرة القرار عن بيانات 2018 و 2019 و2020 | 29 |
| 77 | اختيار خوارزمية بيبز المبسط (Naïve Bayes) عن بيانات 2018 | 30 |
| 77 | ملف السمات عن بيانات 2018 | 31 |
| 78 | احتماليات التعثر وعدم التعثر وفق خصائص السمات لبيانات 2018 | 32 |
| 80 | ملف السمات عن بيانات 2018 و2019 | 33 |
| 82 | احتماليات التعثر وعدم التعثر وفق خصائص السمات لبيانات 2018 و2019 | 34 |
| 84 | ملف السمات عن بيانات 2018 و2019 و2020 | 35 |
| 85 | احتماليات التعثر وعدم التعثر وفق خصائص السمات لبيانات 2018 و2019 و2020 | 36 |

قائمة الملاحق

| رقم الصفحة | الموضوع | رقم الملحق |
|------------|----------------------------|------------|
| 94 | الهيكل التنظيمي لبنك الشام | 1 |

مستخلص الدراسة باللغة العربية

الطالبة: ريم فريد محمد شيخي

العنوان: دور التنقيب في البيانات في اتخاذ قرارات المنح الائتمان دراسة حالة بنك الشام

عام: 2021

الجامعة: الجامعة الافتراضية السورية

مشرف البحث: الدكتور حيدر عبد الله

أدى التطور الهائل في مجال تكنولوجيا المعلومات الى توفر كمية كبيرة من البيانات والتي يمكن أن تساهم بصورة مباشرة في دعم قرارات متخذي القرار إذا ما تم تحليلها بالصورة المطلوبة.

هدفت هذه الدراسة الى تحليل بيانات التمويلات الحالية لدى بنك الشام وذلك من خلال استخدام خوارزميات التنقيب في البيانات، وبناء نموذج تنبؤي للتنبؤ بسلوك سداد المقترضين وتعثراتهم، وباستخدام تقنيات التنقيب في البيانات وبالاستدلال بهذه النماذج تُمكن متخذ القرار من اتخاذ قرار ائتمانية أكثر صحة وأكثر دقة وبالتالي حماية المصرف من الخطر المحتمل والتقليل من تكاليفها، وكما تهدف هذه الدراسة الى التعرف على تقنيات التنقيب في البيانات والى آلية وأهمية تطبيقها في المصارف.

يشمل مجتمع الدراسة العملاء المستفيدين من التسهيلات التجارية والتجزئة خلال الأعوام من 2018 الى 2020 وقد شملت العينة 1987 سجل.

استخدمت الباحثة برنامج SQL server لخلق قاعدة البيانات وبرنامج Microsoft SQL Analysis Server لتطبيق أدوات التنقيب في البيانات لعرض النتائج وتحليلها، وكما تم تطبيق خوارزميات شجرة القرار و ببيز المبسط (Naïve Bayes) .

وقد بينت النتائج التي تم التوصل إليها أن أفضل تقنية للتنبؤ هي شجرة القرار حيث أعطت درجة دقة لبيانات عام 2018 وكانت بنسبة 90.2% وزادت النسبة عند اجراء التنقيب على بيانات 2018 و 2019 وكانت بنسبة 92.7% بينما كانت النسبة 91.9% عند اجراء التنقيب على بيانات الأعوام 2018 و 2019 و 2020 ويعود ذلك لانخفاض عدد العملاء الحاصلين على تسهيلات في عام 2020، بينما أعطت نتائج تنقيب البيانات عند تطبيق خوارزمية ببيز المبسط (Naïve Bayes) بحيث أعطت درجة دقة النموذج 88.8% وذلك عن بيانات 2018 بينما كانت درجة دقة النموذج 92.1% عن بيانات عامي 2018 و 2020 بينما انخفضت درجة دقة النموذج الى 87.9% عند اجراء التنقيب على بيانات الأعوام 2018 و 2019 و 2020.

وباستخدام تقنيات التنقيب في البيانات سيساعد متخذي القرار في البنك من تحليل المخاطر المتعلقة بمنح القروض والتنبؤ بمخاطرها قبل وقوعها، كما سيوفر النموذج الكثير من الوقت والجهد في عملية دراسة القروض وتحليلها.

الكلمات المفتاحية:

المنح الائتماني، متخذي القرار، شجرة القرار، بيزز المبسط (Naïve Bayes)، التنقيب في البيانات.

ABSTRACT

The great development in the field of information technology has led to the availability of a large amount of data that can directly contribute to supporting the decisions of decision makers if they are analyzed in the required manner.

This study aimed to analyze the current financing data of Cham Bank, by using the data mining algorithms, and building a predictive model to predict borrowers' repayment behavior and defaults. By using data mining techniques and by inferring these models, the decision maker could make a better credit decision and can protect the bank from probable risk and reduce its costs. This study aims to identify the techniques of data mining and the mechanism and importance of their application in banks.

The study population includes customers benefiting from commercial and retail facilities during the years from 2018 to 2020, and included 1987 records.

The researcher used the SQL server program to create databases and the Microsoft SQL Server Analysis program to apply data mining tools to display and analyze the results. Decision tree and Naïve Bayes algorithms were also applied.

The results that were reached showed that the best prediction technique is the decision tree, which gave an accuracy degree of the 2018 data and was at a rate of 90.2%, and the percentage increased when the drilling procedure was entered into the 2018 and 2019 data and it was at a rate of 92.7%, while the percentage was 91.9% when conducting an excavation on the data of the years 2018, 2019, and 2020 due to the decrease in the number of clients obtaining facilities in 2020, while the results of data mining when applying Naïve Bayes gave 88.8% accuracy compared to 2018 data, while the model's accuracy was 92.1% compared to yearly data. 2018 and 2020, while the accuracy of the model decreased to 87.9% when excavating the data for the years 2018, 2019, 2020.

By using data mining techniques, it will help the bank's decision makers to analyze the risks related to granting loans and predict their risks before they occur. The model will also save a lot of time and effort in the process of studying and analyzing loans.

Key words:

Credit Grants, Decision Makers, Decision Tree, Naïve Bayes, Data Mining.

الفصل الأول: الإطار المنهجي للدراسة

مقدمة:

يُعتبر الائتمان المصرفي من أهم الوظائف الرئيسية التي تقوم بها المؤسسات المالية (شاهين، 2010)¹، إذ يلعب دوراً هاماً في تحقيق الربحية لها، وذلك لكونه من أهم المصادر لإشباع الحاجات التمويلية لقطاعات النشاطات الاقتصادية المختلفة، وبالتالي فإن ارتفاع نسبة القروض في ميزانيات المصارف يؤدي الى ارتفاع الفوائد والعمولات التي تعتبر كمصدر للإيرادات، ويؤدي ارتفاع نسبة القروض الممنوحة الى ارتفاع المخاطر المرتبطة بها نتيجة احتمال عدم سدادها (زميت، 2005)².

وكما ينشأ الائتمان المصرفي من خلال قيام البنك بتوظيف أموال المودعين عن طريق منحها لطالبي الائتمان، يعتبر البنك ملتزم بالوفاء بأموال المودعين حال انتهاء آجالها أو عند طلبها، فإن تعثر الحاصلين على الائتمان يعرض البنك إلى خسائر مالية تؤثر على إيراداته وعلى مركزه المالي (كراسنة، 2006)³.

بدأت المصارف ببناء سياسات ائتمانية لتسترشد بها المستويات الإدارية المختلفة لأخذ القرار في الطلبات الائتمانية المقدمة حسب شروط وقيود معينة يفرضها البنك على عملائه، كما يعتبر اتخاذ هذه القرارات الائتمانية هو جوهر العملية المصرفية ونجاحها مرتبط بشكل كبير بحجم المعلومات وبنوعيتها المتوافرة لدى المصرف.

وبازدياد البيانات الموجودة والمخزنة بقواعد البيانات (Database)، وانتشار مستودع البيانات (Data warehouse)، أصبح من الضروري البحث عن أدوات تمتاز بالقوة لتحليل هذه البيانات واستخراج المعلومات والمعارف منها.

ظهر في أواخر الثمانينات ما يسمى بالتنقيب في البيانات (Data mining) الذي أثبت وجوده كأحد الحلول الناجحة لتحليل كميات ضخمة من البيانات، وذلك بتحويلها من مجرد معلومات متراكمة وغير مفهومة (بيانات) إلى معلومات قيّمة يمكن استغلالها والاستفادة منها (الكاشف، 2000)⁴.

وباعتبار التنقيب في البيانات من التقنيات الحديثة التي فرضت نفسها بقوة في عصر المعلوماتية، واستخدامها يوفر للشركات والمنظمات في جميع المجالات القدرة على الاستكشاف والتركيز على أهم المعلومات

¹ شاهين، علي عبد الله. (2010) "مدخل عملي لقياس مخاطر الائتمان المصرفي في البنوك التجارية - دراسة تحليلية تطبيقية". مكتبة الجامعة الإسلامية.

² زميت، محمد (2005)، "النظام المصرفي الجزائري في مواجهة تحديات العولمة المالية". رسالة ماجستير، كلية الاقتصاد، جامعة يوسف بن خدة.

³ كراسنة، ابراهيم. (2006). أطر أساسية و معاصرة في الرقابة على البنوك و إدارة المخاطر. أبو ظبي: معهد السياسات الاقتصادية.

⁴ الكاشف، محمود يوسف (2000)، "مدخل مقترح لتطوير دور المعلومات المحاسبية في إطار المفهوم المتكامل للجودة الشاملة"، مجلة الإدارة العامة، مجلد 94، العدد 3، أكتوبر.

الموجودة في قواعد البيانات، كما تُركز تقنيات التنقيب على بناء التنبؤات المستقبلية واستكشاف السلوك والاتجاهات مما يسمح باتخاذ قرارات سليمة.

ومن هنا يأتي دور التنقيب في البيانات في المؤسسات المصرفية من خلال مساعدتها في الاستفادة من البيانات الهائلة والمتوافرة لديها والقيام بتحليلها مما يؤدي إلى الحد من المخاطر المحتملة، كالمخاطر الناجمة عن عدم سداد العملاء للالتزاماتهم، والتنبؤ بالمقترضين الجيدين بهم واستبعاد غير الجيدين، ومساعدة صانعي القرار في اتخاذ قرارات ائتمانية سليمة.

أولاً-مشكلة البحث:

تعد خسائر الائتمان أمر محتمل حدوثه كنتيجة لعملية الإقراض، وكل مصرف يتحمل درجة من الخطر عند قيامه بعملية منح ائتماني، كما يحقق المصرف بعض الخسائر عندما يفشل في استرداد قرضه، وإن خطر الائتمان يعني تخلف العملاء عن الدفع أو عجزهم عن السداد أو خسارة كُلية أو جزئية لأي مبلغ مُقرض إلى الطرف المقابل، ولهذا فإن عملية المنح الائتماني بحاجة إلى وجود سياسة ائتمانية جيدة بالإضافة إلى التنبؤ بسداد القروض لحماية أصول المصرف من الضياع.

ولقيام المصرف بعملية التنبؤ بالسداد يجب عليها أن تنظر بعمق إلى بياناتها التاريخية بهدف تحليلها وذلك بغرض التخفيف من الخسائر الائتمانية المحتملة ودعم قراراتها.

ومن خلال عمل الباحثة في بنك الشام واطلاعها على البرمجيات والتقنيات المتوافرة لدى البنك فلاحظت غياب تقنيات دعم القرار والتي يُمكن استخدامها في التنبؤ بسلوك المقترضين والتخفيف من الخسائر المحتملة من خلال دعم قراراتهم الائتمانية.

وبالتالي تتبلور مشكلة البحث بما يلي:

ما هو دور التنقيب في البيانات في اتخاذ قرارات المنح الائتمان دراسة حالة بنك الشام؟
ويتفرع عنه التساؤلات التالية :

- 1- إلى أي حد يجري العمل على الاستفادة من البيانات المتوفرة لدعم عملية اتخاذ القرار الائتماني؟
- 2- هل يوجد آلية لدى المصرف للتنبؤ بتعثر العملاء؟
- 3- هل لخوارزمية شجرة القرار دور في دعم القرار الائتماني؟
- 4- هل لخوارزمية بيبز المبسط (Naïve Bayes) دور في دعم القرار الائتماني؟
- 5- ماهي أفضل أساليب التنقيب التي يمكن استخدامها لتقييم مخاطر التسهيلات الائتمانية؟

ثانياً- أهداف البحث:

يهدف البحث لبيان دور التنقيب في البيانات في عملية اتخاذ قرار المنح الائتماني، وذلك من خلال تحقيق مجموعة من الأهداف والتي تتمثل فيما يلي:

- 1- تطبيق مفاهيم التنقيب في البيانات.
- 2- استخدام خوارزمية شجرة القرار للمساعدة في دعم القرار الائتماني.
- 3- استخدام بيزز المبسط (Naïve Bayes) للمساعدة في دعم القرار الائتماني .
- 4- بناء نموذج تنبؤي من خلال خوارزميات شجرة القرار و بيزز المبسط (Naïve Bayes) من خلال البيانات التاريخية الذي يساعد في دعم قرارات المنح الائتماني.
- 5- التوصل إلى أفضل أساليب التنقيب التي يمكن استخدامها لتقييم مخاطر التسهيلات الائتمانية من خلال مقارنة نتائج النماذج التنبؤية الخاصة بشجرة القرار و بيزز المبسط (Naïve Bayes).

ثالثاً- أهمية البحث:

تأتي أهمية هذا البحث من أهمية القرارات الائتمانية في المصارف من جهة، وأهمية استخدام أدوات التنقيب في البيانات من جهة أخرى، وكما جاءت أهمية هذا البحث نتيجة لقلّة الدراسات التي تناولت هذا الموضوع نظراً لحدائته، وخاصةً فيما يتعلق بالجانب الائتماني، ويمكن تلخيص أهمية البحث الى أهمية علمية وأهمية تطبيقية:

1-أهمية علمية:

- 1- بسبب الكمية الهائلة من البيانات التي يجري تخزينها بشكل مستمر في قواعد البيانات أصبح التنقيب في البيانات أحد الحلول الفعّالة للمساعدة في الاستفادة من هذه البيانات بصورة جيدة.

2-أهمية تطبيقية:

- 1- رفع كفاءة وجودة الخدمات المقدمة لدائني المصرف.
- 2- نظراً لتشابه الخدمات والمنتجات المصرفية وتشابه أسلوب تطويرها في المصارف السورية، فمن الممكن استخدام هذا البحث من قبل المصارف الأخرى في سورية والذي بدوره يسهم في دعم قراراتهم الائتمانية بالإضافة إلى التقليل من المخاطر الائتمانية المحتملة.
- 3- المساهمة في تقييم مخاطر الائتمان باستخدام التنقيب في البيانات وللكشف عن قواعد معرفة جديدة والتي كانت غير معروفة أو من الصعب معرفتها دون استخدام خوارزميات التنقيب في البيانات.

رابعاً-منهجية البحث:

تتبع الدراسة المنهج الوصفي التحليلي والتطبيقي لوصف وتحليل البيانات. وسيتم الاعتماد في الدراسة على استراتيجيات دراسة الحالة العملية حيث أن الدراسة ستركز على شركة واحدة وهي شركة بنك الشام، بحيث يتم جمع بيانات تفاصيل التمويلات الائتمانية (شركات وأفراد). ولتحقيق أهداف البحث سيتم اعتماد الخطوات المنهجية التالية:

- 1- جمع بيانات التمويلات من بنك الشام.
- 2- تنقية البيانات (Data cleaning).
- 3- بناء قاعدة بيانات تتوافق مع البيانات التي تم جمعها.
- 4- توريد البيانات الى قاعدة البيانات.
- 5- اختيار البرامج المناسبة لعملية تنقيب البيانات.
- 6- تحديد الواصفة المراد التنبؤ بها وتحديد الواصفات الغير مهمة.
- 7- اختيار الخوارزميات المراد استخدامها.
- 8- استكشاف المعارف والقواعد الجديدة من خلال كل خوارزمية على حدى.
- 9- مناقشة وتقييم النتائج.

خامساً-حدود البحث:

- الحدود الزمانية: سيتم اجراء هذه الدراسة على بيانات الفترة الزمنية 2018-2020 م.
- الحدود المكانية: بنك الشام – سورية.

سادساً-الدراسات السابقة:

هناك العديد من الدراسات ذات صلة بالموضوع والتي أمكن الحصول عليها، وفيما يلي أمثلة لبعض هذه الدراسات:

الدراسات العربية:

- 1- دراسة (عبد المجيد،2018): " استخدام تقنيات تنقيب البيانات في تحليل ومعرفة تفاصيل عمليات التمويل الأصغر، دراسة حالة بنك فيصل الإسلامي السوداني"⁵

⁵ عبد المجيد، أمل محمد عز الدين.(2018): " استخدام تقنيات تنقيب البيانات في تحليل ومعرفة تفاصيل عمليات التمويل الأصغر، دراسة حالة بنك فيصل الإسلامي السوداني". رسالة ماجستير غير منشورة، كلية الدراسات العليا، جامعة النيلين.

هدفت هذه الدراسة إلى استخدام تقنيات التنقيب في البيانات في حل المشاكل واتخاذ القرارات في عمليات التمويل الأصغر وذلك من أجل تفادي التعثر من خلال معرفة تفاصيل عمليات التمويل الناجحة. وتمثلت مشكلة البحث في صعوبة اتخاذ قرار منح التمويل من قبل البنك، ومعرفة المشاريع الناجحة بالنسبة للمؤسسة المقدمة للتمويل، وفشل نفس المشاريع مع عدم معرفة السبب، ووجود بيانات ضخمة لم يستفاد منها. وتم اعتماد المنهج الوصفي التحليلي والتطبيقي لوصف وتحليل البيانات، وتقوم على جمع البيانات للاطلاع على مدى نجاح التمويل الأصغر وتفاصيل المشاريع الناجحة والاستفادة من البيانات للتمويل المستقبلي. وجاءت نتيجة الدراسة بعد تطبيق كل من خوارزميتي C4.5 في التصنيف و خوارزمية (K-means) إلى في تقنية العناقيد إلى أن أكثر عمليات التمويل غير متعثرة وتفاصيل عمليات التمويل الناجحة هي من حيث الولاية أكثرها في ولاية الخرطوم، والقطاع هو قطاع التجارة، ومن حيث السلع فهي المواد الغذائية، وغالبية عمليات التمويل بصيغة المرابحة، والضامن لعمليات التمويل هو شيك الضمان، والمبلغ المصدق للتمويل هو 20,000 ألف جنيه سوداني ثم يليه 30,000 ألف جنيه سوداني، وأن غالبية طالبي التمويل الأصغر بالبنك هم من الرجال.

2- دراسة (فقيري، 2018): "تطبيق تقنية تعلم المكنان للتنبؤ بمخاطر التمويل البنكي – دراسة حالة بنك العمال الوطني"

هدفت هذه الدراسة الى استخدام وتطبيق خوارزميات تعلم المكنان في بناء النموذج المناسب للتنبؤ باكتشاف حدوث خطر يهدد البنك بالتعرف على الخصائص التي تشير الى الخطر المحتمل حدوثه، وتتمثل المخاطر في عدم ايفاء العملاء بالسداد، وبالتالي فإن استخدام تقنيات التنقيب وبلاستدلال بهذه النماذج من خلال البيانات المتاحة، تمكن متخذي القرار من التعرف على الخطر في فترة زمنية معقولة قبل حدوث الخطر وبالتالي حماية المصرف من هذا الخطر المستقبلي.

وأظهرت النتائج أنه ومن خلال استخدام النظم الذكية يمكن للبنك التنبؤ بالمخاطر الائتمانية مما يمكن البنك من وضع الاحتياطات اللازمة لحل هذه المشكلة. لقياس دقة اكتشاف المخاطر تم استخدام مصفوفة الارتباك والتي تحتوي معيار الصواب والخطأ وعلى صنف الموجب والسالب، تم استخدام شجرة القرار في عملية التصنيف واستخدمت لاكتساب المعلومات وتقسيم الشجرة للتنبؤ بحدوث الخطر كما أن خوارزمية شجرة القرار حققت نسبة دقة 100% في اكتشاف المخاطر مقارنة بخوارزمية المستقبل متعدد الطبقات والتي كانت نسبة دقة نتائجها 97.7%، من خلال بيانات رقمية وحرفية بحجم 3847 عمل.

6 حسن، وضاح عبدالله، فقيري، سلام عثمان. (2018): "تطبيق تقنية تعلم المكنان للتنبؤ بمخاطر التمويل البنكي – دراسة حالة بنك العمال الوطني". بحث منشور. كلية علوم الحاسوب وتقانة المعلومات، جامعة الزعيم الأزهرى.

3- دراسة (الخليفة، 2017): "استخدام تنقيب البيانات في القطاع المصرفي-دراسة حالة في المصارف السودانية"⁷

هدفت هذه الدراسة الى القاء الضوء على أهمية تنقيب البيانات في القطاع المصرفي من خلال التركيز على عملية التنبؤ بنمط العملاء وتعاملهم مع العمليات الاستثمارية، ودراسة مشكلة تراكم البيانات ومحاولة إيجاد الحل الأمثل، وتحسين مستوى استخدام المعلومات من قبل الإدارات من اجل زيادة أداء، بالإضافة الى بناء نماذج تنقيب البيانات التي يمكن أن تستخدم في التنبؤ بالمقترضين الموثوق بهم واستبعاد غير الموثوق بهم لتقليل تكاليفهم.

وتمثلت مشكلة البحث بوجود عدد كبير من العملاء في المصرف ووجود كم هائل من البيانات عنهم وعدم توافر تقنية تمكن المصرف للاستفادة من هذه البيانات ومساعدته في دعم قراراته واتخاذها بشكل سريع. ويعتمد منهج البحث على المنهج الوصفي والتحليلي من أجل استخلاص المؤشرات الأساسية التي تدعم عملية اتخاذ القرار.

وجاءت نتيجة الدراسة بأن:

- 1- تنقيب البيانات والتنبؤ بصورة خاصة يساعد في تقليل نسبة التكاليف والوقت والجهد.
- 2- الطرق التنقيبية ناجحة جداً في دعم القرار، وتصنيف عملاء البنوك والنمط الجامع بين العملاء في مختلف مجالات الخدمات المقدمة لدى البنوك.
- 3- برنامج التنبؤ بالأنماط يؤدي الى مواصلة التقدم في الجودة الشاملة للنظام المصرفي وتقديم خدمات أفضل.
- 4- الاستمرار في نظام التنبؤ في ظل نظام واضح، متطور، متكامل، يحافظ على السرعة مصحوبة بالدقة في انجاز الربح المؤكد.

4- دراسة (نصر الله، 2017): " التنبؤ بسداد قروض المصارف باستخدام تنقيب البيانات -دراسة الحالة) بالتطبيق في بنك فيصل الإسلامي"⁸

هدفت هذه الدراسة إلى تطبيق مفاهيم التنقيب في البيانات في بنك فيصل الإسلامي، للتنبؤ بمعرفة العملاء الصالحين في سداد القروض الاستثمارية باختلاف الجنس ومجال العمل والمستوى الوظيفي والمؤهل العلمي والفئة العمرية والحالة الاجتماعية للعملاء.

⁷ الخليفة، شبكية هاشم الطيب.(2017): " استخدام تنقيب البيانات في القطاع المصرفي-دراسة حالة في المصارف السودانية". رسالة ماجستير غير منشورة، كلية الدراسات العليا، جامعة النيلين.

⁸ نصرالله، هبة مبارك موسي.(2017): " التنبؤ بسداد قروض المصارف باستخدام تنقيب البيانات -دراسة الحالة بالتطبيق في بنك فيصل الإسلامي)". رسالة ماجستير غير منشورة، كلية الدراسات العليا، جامعة النيلين.

وتمثلت مشكلة البحث بوجود كم هائل من البيانات والمعلومات المحاسبية وغير المحاسبية المخزنة والتي تعد من أهم أشكال المعرفة الواجب البحث فيها لضمان ديمومة وكفاءة العمليات المصرفية المنفذة والمعمول بها لدى البنوك لأهمية دورها في توفير المعلومات اللازمة لتلبية احتياجات المستخدم الداخلي والخارجي للبنوك، ويتطلب البحث في كيفية استخدام الفعال لهذه البيانات لغايات التعرف على مختلف المتغيرات فيما بينها وكيفية تأثيرها على عمل البنوك وعملياتها المحاسبية والمصرفية.

وتم اعتماد المنهج الوصفي والتطبيقي من خلال استخدام خوارزميات التنقيب في البيانات واتخاذ القرار الأمثل.

وجاءت نتيجة الدراسة بعد اجراء العديد من التجارب على برنامج (Orange&Weka) ودراسة بيانات العملاء بأخذ في الاعتبار العمر والمستوى العلمي ومستوى الدخل والحالة الاجتماعية وتوفر السكن للعميل ودراسة الارتباط بينهما مما يساعد في تسهيل اتخاذ القرار للبنك في معرفة العملاء الأنسب لوضع استراتيجيات تجذب العملاء بصورة أكبر، النتائج التي توصل اليها الباحث:

- 1- عدد عملاء البنك الذي تم اختيارهم للعينة 2000 عميل قد تم تصنيف 1808 عميل للبنك في منحهم القروض و191 عميل تم استبعادهم في أخذ القرض.
- 2- الفئة العمرية التي تأخذ القروض الاستثمارية هي من 20-60 سنة.
- 3- المتزوجون هم من أكثر الفئات الأخذة للقروض الاستثمارية.
- 4- الفئة المتعلمة أكثر اقبال على أخذ القروض الاستثمارية من الفئة غير المتعلمة.

5- دراسة (حسن، 2015): " تعدين بيانات التمويل الأصغر باستخدام تقنيات التصنيف والعنقدة"⁹

يهدف البحث الى استخدام طرق ذكية لاستخلاص أنماط البيانات واستخراج نماذج مفيدة بقدر الإمكان، وتمثلت مشكلة البحث بوجود كم هائل من البيانات مخزن لدى المصرف، بالإضافة لحل مشكلة فشل المقترضين بعدم سداد مديونياتهم.

وتم اعتماد المنهج الوصفي والتطبيقي من خلال استخدام خوارزمية شجرة القرار لمعرفة حالة التمويل، من خلال دراسة حالة مؤسسة الشباب للتمويل الأصغر.

وجاءت نتائج الدراسة بمايلي :

- 1- تم معرفة تفاصيل القروض الخطرة والناجحة وذلك عن طريق نتائج شجرة القرار.
- 2- أكثر أنواع المنتجات نجاحاً هو البيع التجاري بنسبة 19%، وفترة سداد 13 شهر، يجب العمل بهذه النتيجة لتفادي التمويل الخطر.

⁹ حسن، هالة حسن محمود.(2015): " تعدين بيانات التمويل الأصغر باستخدام تقنيات التصنيف والعنقدة ". رسالة ماجستير غير منشورة، كلية الدراسات العليا، جامعة النيلين.

3- أكثر أنواع المنتجات فشلاً في مجال الزراعة فبالثالي هي من المنتجات التي يجب تجنبها لأنها تمثل خطر على التمويل.

4- أعلى نسبة خطر تتمثل في المشاريع الخدمية الأخرى وهي نسبة كبيرة جداً، ومشروع صيانة عربات الى حد ما.

6- دراسة (الدوري، أحمد، 2007): " دور تنقيب البيانات في زيادة أداء المنظمة-دراسة تحليلية في المصرف الصناعي العراقي" ¹⁰

تتلخص مشكلة البحث في ثلاثة نقاط أساسية هي: وجود عدد كبير من المقترضين في المصرف الصناعي ووجود كم هائل من البيانات عنهم وعدم وجود تقنية تمكن المصرف من اتخاذ قرار سريع لإقراضهم. يقدم البحث نموذج يسهل عملية تنقيب البيانات في المصرف حيث ينتبأ بالمقترضين ويصنفهم حسب مخاطر التسديد لكي يتمكن المصرف من التركيز على المقترضين الموثوق بهم فقط سعياً منه لزيادة الأداء.

وتم اعتماد المنهج الوصفي والتطبيقي من خلال استخدام خوارزمية شجرة القرار وخوارزمية التجميع. وجاءت نتيجة الدراسة:

- 1- بيانات شجرة القرار تبين عدد حالات الخطر السيئة 97 حالة بنسبة 35% وعدد حالات الخطر الجيدة 180 حالة وبنسبة احتمالية 64.64% ولا توجد حالات مفقودة.
- 2- عدد الحالات السيئة للقروض تصل الى 28 حالة خطر سيئة مقابل 2 حالة خطر جيدة فقط مع نسب احتمالية كل منها. وتنحصر مخاطر القروض السيئة في فروع أربيل والسليمانية والمستثمرة في الصناعات الانشائية والميكانيكية ومحطات تعبئة الوقود والنجارة وتكون مستويات القروض متوسطة وعالية والمقترضين هم أفراد من كلا الجنسين ذكور واناث.
- 3- مخاطر القروض جيدة لفروع المصرف في نينوى وبغداد والبصرة والصناعات هي الجلود والتلج والملح والغذائية والنجارة والطحين والبلاستيك والكيمياوية والجنس هو ذكر ومستوى القرض واطى وأصحاب القروض هم أفراد منهم متزوجون واخرون غير متزوجين.

¹⁰ الدوري، زكريا مطلق؛ أحمد، داليا عبد الحسين (2007)، " دور تنقيب البيانات في زيادة أداء المنظمة-دراسة تحليلية في المصرف الصناعي العراقي"، مجلة العلوم الاقتصادية والإدارية، العدد 48، العراق.

"Loan default prediction using decision trees and random forest: A comparative study"

"التنبؤ باحتمالية تعثر القرض باستخدام أشجار القرار و الغابات العشوائية: دراسة مقارنة"

هدفت هذه الدراسة إلى استكشاف وبناء خوارزميات التنقيب في البيانات لتحديد ما إذا كان الشخص بالنظر إلى سمات معينة لديه احتمال كبير للتخلف عن سداد القرض. يمكن استخدام هذا النوع من النموذج لتحديد سمات مالية معينة للمقترضين المستقبليين والتي يمكن أن يكون لها القدرة على التخلف عن السداد وعدم سداد قروضهم في الوقت المحدد.

وجاءت نتائج الدراسة بأن Random Forest Classifier يقدم دقة تصل إلى 80٪ بينما زودت طريقة شجرة القرار بدقة 73٪. ومن ثم يبدو أن نموذج الغابة العشوائية هو خيار أفضل لمثل هذا النوع من البيانات.

وكما تبين من خلال الدراسة على إمكانية تطبيق النموذج التنبؤي للتنبؤ بقرارات المنح الائتماني ولكن بعد تدريب بيانات أكثر تعود لبيانات تاريخية لمساعدة النموذج على التنبؤ بدقة بالمقترضين المؤهلين .

"Loan default forecasting using Data mining"

" التنبؤ بتعثر القرض باستخدام التنقيب في البيانات"

هدفت هذه الدراسة الى تقييم العملاء المتعثرين عن السداد وذلك باستخدام تقنيات التنقيب في البيانات، وجاءت أهمية الدراسة حيث تعتبر عملية تقييم العملاء من العمليات ذات الأهمية في المصارف والتي يجب على المصارف بتنفيذها لمساعدتهم في اتخاذ قرارات منح ائتماني أكثر صواباً وأقل خطراً. ونتيجة لقدرة هذه التقييمات في مساعدة المؤسسات المالية على تقليل خسائرها وزيادة عدد القروض الجيدة واتخاذ قرارات سليمة، فأصبح من الضروري بناء نموذج يأخذ في الاعتبار الجوانب المختلفة لمقدمي

¹¹ Madaan,Mehul.Kumar,Aniket.Keshri,Chirag.Jain,Rachna.and Nagrath,Preeti(2021) " Loan default prediction using decision trees and random forest: A comparative study",A thesis, Department of Electronics and Communication Engineering, GGSIP University.

¹² Patel,Bhoomi.Patil,Harshal.Hembram,Jovita.and Jaswal,shree(2020) "Loan default forecasting using data mining",International conference for emerging technology(INCET), Jun5-7,2020.

الطلبات الائتمانية. قام الباحثين باستخدام خوارزميات التنقيب في البيانات للتنبؤ بالمقترضين المتعثرين بناء على مجموعة من البيانات التاريخية. وجاءت نتائج البحث بأنه تم الحصول على أفضل النتائج باستخدام خوارزميات الانحدار اللوجستي، Random forest، Gradient Boosting، Catboost classifier. وكما أكدت الدراسة على إمكانية تطبيق هذا النموذج لاتخاذ قرارات أفضل بشأن طالبي القروض الائتمانية وتخفيض من الخسائر الائتمانية المحتملة.

3- دراسة (Alemu, Teninet Belay, 2019)¹³

"Predicting bank credit risk using Data mining Technique: the case of bank of Abyssinia"

"التنبؤ بمخاطر الائتمان المصرفي باستخدام تقنية التنقيب عن البيانات: حالة بنك الحبشة"

هدفت الدراسة الى تطبيق تقنيات التنقيب في البيانات للتنبؤ بمخاطر الائتمان المصرفي في بنك Abyssinia ، حيث تمثلت مشكلة البحث في حاجة المصرف الى توسيع استثماراته عن طريق تقديم منح ائتماني لعملاء مختلفين، ولكن هذا المنح قد يعرض المصرف الى خطر عدم السداد. ولأهمية تحديد العملاء المتعثرين في السداد فكان لا بد من بناء نموذج تنبؤي يساعد في تحديد درجة المخاطرة عند اتخاذ قرار المنح الائتماني. قام الباحث ببناء نموذج تنبؤي من خلال تحليل البيانات التاريخية للعملاء الحاصلين على قروض ائتمانية وذلك من خلال تقنيات التنقيب في البيانات .

وجاءت نتيجة الدراسة بعد تطبيق ثلاثة خوارزميات لاستخراج المعلومات وهي:

J48- Decision tree- JRip Classification Rule- Naive Bayes

وتبين بأن شجرة القرار J48 هي أفضل نموذج لأغراض التنبؤ والتي سجلت أعلى نسبة دقة 97.52% .

4- دراسة (Feyissa , 2018)¹⁴

" The use of Data mining to predict the loan repayment risk: the case of Oromia credit and saving share company"

"استخدام التنقيب عن البيانات للتنبؤ بمخاطر سداد القرض: حالة شركة أوروميا للائتمان والادخار"

¹³ Alemu, Teninet Belay (2019) "Predicting bank credit risk using data mining Technique: the case of bank of Abyssinia", A thesis, faculty of informatics, ST. Mary's university.

¹⁴ Feyssa, Ketema (2018) " The use of data mining to predict the loan repayment risk: the case of Oromia credit and saving share company ", A thesis, college of natural science, Addis Ababa University.

هدفت هذه الدراسة الى تطبيق خوارزميات التنقيب في البيانات لإعداد نموذج التنبؤ بسداد القروض في شركة أوروميا للانتمان والادخار من أجل تحسين عملية اتخاذ القرار، حيث تمثلت مشكلة البحث بأن الإقراض في اثيوبيا أمراً مثيراً للجدل وصعباً من ناحية ومن ناحية أخرى نقص بعدد القروض الائتمانية الممنوحة نتيجة للمعايير العالية التي تقوم بفرضها المؤسسات التمويلية على العملاء، وذلك نتيجة لتخوف المؤسسات المالية من الخسائر التي من المحتمل أن تتعرض لها، والإقراض بطبيعته يتطلب أن يثق المقرض بالمقترض لسداد القرض في تاريخه المحدد من جهة، ومن جهة أخرى يتطلب من المقرض أن يتوفر لديه أدوات ووسائل تساعده في التمييز بين العملاء الجيدين والجديرين بالثقة عن غيرهم. قام البحث بإعداد نموذج باستخدام تقنيات التنقيب في البيانات لمساعدة متخذي القرار باتخاذ قرارات ائتمانية سليمة وتعزيز ثقته بطالب الائتمان. وجاءت نتيجة الدراسة بأن أكثر النماذج فاعليه هو شجرة القرار J48 بدرجة دقة 98,89 %، وكما أكدت النتائج على ان القواعد التي تم انشاؤها على إمكانية تطبيقها في توقع سداد قرض العميل في المؤسسات المالية وقدرتها على تخفيف من الخسائر الائتمان المحتملة.

5- دراسة (Gahlaut , et . al 2017)¹⁵

" Prediction analysis of risky credit using Data mining classification models"

"تحليل التنبؤ بالائتمان الخطر باستخدام نماذج تصنيف التنقيب في البيانات"

هدفت هذه الدراسة الى التحقق فيما إذا كانت تقنيات التنقيب في البيانات مفيدة للتنبؤ بدرجة ائتمان العميل وتصنيفه، وذلك من أجل التغلب على مخاطر المنح الائتماني للعملاء في المستقبل.

قام الباحثون بالاعتماد على بيانات تاريخية من أجل بناء النماذج التنبؤية والتي تمكن البنك في استخدامها للحصول على نتائج أفضل لنظام الائتمان العام، والذي يساعد على اتخاذ القرار الصحيح فيما يتعلق بالموافقة أو رفض طلب الخاص بالعميل بناءً على خلفيته العائلية، المالية، المهنية، وحالته الاجتماعية، بالإضافة الى عوامل أخرى.

اعتمد الباحثون في دراستهم على استخدام خوارزميات التنقيب في البيانات مثل شجرة القرار، الانحدار الخطي، Random forest، الشبكة العصبية لبناء نماذج للتنبؤ بالقروض وتصنيفاتها.

تبين نتائج البحث أن أفضل خوارزمية لتصنيف الائتمان والتحوط من المخاطر المحتملة هي خوارزمية Random forest لكونها تتمتع بالدقة العالية على الرغم من كونها الأقل سرعة في وقت التشغيل الا أنها تعطي أفضل أداء وأعلى دقة.

¹⁵ Gahlaut, Archana. Tushar. and singh, Prince Kumar (2017) " Prediction analysis of risky credit using Data mining classification models", 8th International Conference on Computing, Communication and Networking Technologies (ICCCNT), 5-3 July 2017.

"Developing prediction model of loan risk in banks using Data mining"

"تطوير نموذج التنبؤ بمخاطر القروض في البنوك باستخدام التنقيب عن البيانات"

هدفت هذه الدراسة الى تطوير نموذج لتصنيف مخاطر القروض في القطاع المصرفي باستخدام تقنيات التنقيب في البيانات.

والذي جاء نتيجة الى حاجة المصارف الى تحليل المخاطر المحتملة من القروض المصرفية، بالإضافة الى تزايد نمو المعاملات في القطاع المصرفي وتوافر كميات ضخمة من البيانات والتي تمثل سلوك العملاء دون الاستفادة منها، وكما يعد التنقيب في البيانات أحد المجالات الحيوية لاستخراج معلومات ومعارف جديدة من الكميات المتركمة من البيانات.

قام الباحثون ببناء نموذج باستخدام بيانات القطاع المصرفي للتنبؤ بحالة القرض، حيث تم استخراج ثلاثة خوارزميات لبناء هذا النموذج وهي: J48-BayesNet-NaiveBaye باستخدام تطبيق WEKA .

وبعد مقارنة نتائج التجارب الثلاثة تبين أن أفضل الخوارزميات لتصنيف القروض هي J48 لأنها تتميز بالدقة العالية وانخفاض متوسط الخطأ، وكما جاءت النتائج بالتأكيد على قدرة النموذج على مساعدة متخذي القرار في قبول طلب المنح الائتماني أو رفضه والتوقع فيما إذا كان القرض سيؤدي الى مخاطر وخسائر للبنك.

الفرق بين الدراسة الحالية والدراسات السابقة:

تختلف هذه الدراسة عن الدراسات السابقة في الحدود المكانية حيث سيتم الاعتماد لأول مرة التطرق لدراسة دور التنقيب في البيانات في اتخاذ قرارات المنح الائتماني ضمن بنك الشام في سوريا، وتختلف من حيث البرامج المستخدمة لتطبيق خوارزميات التنقيب في البيانات ففي الدراسات السابقة تم الاعتماد على برنامج WEKA و ORANGE أما في هذا البحث سيتم تطبيق الخوارزميات على برنامج Microsoft SQL Analysis Server والذي يميزه بالتوافقية العالية باعتبار أن كافة البرامج المستخدمة هي من إنتاج شركة Microsoft، وكما تختلف من حيث تحليل البيانات وبناء النماذج حيث سيتم بناء عدة نماذج ومقارنتها مع بيانات واقعية للتحقق من درجة دقة النماذج.

¹⁶ Hamid,Aboobyda Jafar.and Ahmed,tarig Mohammed.(2016)" Developing prediction model of loan risk in banks using Data mining", **Machine Learning and Applications: An International Journal (MLAIJ)**, Vol.3, No.1, March 2016.

الفصل الثاني: الإطار النظري للبحث

أولاً: الائتمان المصرفي:

أهمية المصارف تأتي من الأثر الهام في تحفيزها للعرض الكلي للنقود، فهي لا تكتفي بقبول الودائع بل تقوم بخلقها أيضاً، ويقصد بالنظام المصرفي مجموعة المؤسسات التي تتعامل بالائتمان ، حيث تأتي معظم إيرادات البنوك من الأنشطة الائتمانية وفوائدها، والاهتمام بالوظيفة الائتمانية هو من واجبات المصرف على اعتبار أن المركز المالي لأي مصرف يتأثر بمتغيرات وعناصر كثيرة، ومحفظة القروض بشكل خاص تحتل موقعاً هاماً ضمن بنود المركز المالي، وتتبع أهمية الإقراض في العمل المصرفي من كونه وسيلة لجمع المدخرات لإعادة ضخها في عروق النظام الاقتصادي بصور عديدة كإقراض النقود وفتح الاعتمادات وإصدار خطابات الضمان وخصم الأوراق التجارية وغير ذلك ، فسلامة محفظة القروض يؤدي الى تحقيق عوائد مرتفعة للمصرف عند أقل مستويات ممكنة من المخاطر المصاحبة لقرارات منح الائتمان.

1- مفهوم الائتمان المصرفي:

يعرف الائتمان المصرفي على أنه الثقة التي يوليها البنك لعميله، بحيث يضع تحت تصرفه مبلغاً من المال لاستعماله في غرض محدد خلال فترة محددة مقابل حصول البنك على عائد مالي متفق عليه مع تقديم العميل لضمان يمكن البنك من استرداد أمواله إذا ما توقف العميل عن السداد (بن عمر، 2011) ¹⁷.

إنّ أول أشكال العمل المصرفي بدأ بقبول الودائع دون إعطاء أصحابها حق بالفائدة، بل أنه كان يترتب عليه في بعض الأحيان دفع جزء منها لمن أودعت لديه هذه الممتلكات لقاء حراستها والمحافظة عليها، من ثم أخذت مؤسسات الإيداع بممارسة عمليات الإقراض لقاء فوائد وضمانات تختلف باختلاف طبيعة العمليات الإقراضية (الزبيدي، 2001، ص104) ¹⁸، بعد ذلك ومع تطور العمل المصرفي وتراكم الودائع لدى المؤسسات التي تمارس العمليات المصرفية لاحظت هذه المؤسسات أن قسماً من المودعين يتركون ودائعهم فترة طويلة دون استخدامها، ففكروا باستخدام جزء من هذه الودائع وتسليمها للمحتاجين مقابل فائدة (السيسي، 2004، ص20) ¹⁹.

وهكذا بعد أن كان المودع يدفع عمولة إيداع أصبح يتلقى فائدة على ودائعه، كما انتقل العمل المصرفي من مهمة قبول الودائع الى ممارسة عمليات الإقراض والتسليف، ليصبح الركن الأساسي لأعمال البنوك الحديثة

¹⁷ بن عمر، خالد. (2011): "دراسة النماذج الحديثة لقياس مخاطر الائتمان لدى البنوك التجارية". رسالة دكتوراه غير منشورة، كلية العلوم الاقتصادية والتجارية وعلوم التسيير، جامعة بومرداس.

¹⁸ الزبيدي، محمود. (2001). إدارة الائتمان المصرفي والتحليل الائتماني. الطبعة 1، عمان: دار الوراق للنشر والتوزيع. ص 104.

¹⁹ السيسي، صلاح الدين. (2004). قضايا مصرفية معاصرة في الائتمان-الضمونات المصرفية. الاعتمادات المستندية. القاهرة: دار الفكر العربي. ص 30.

هو قبول الودائع والمدخرات من جهة وتقديم التسهيلات الائتمانية والخدمات المصرفية المتعددة الأشكال من جهة أخرى.

2- أهمية الائتمان المصرفي:

إن عملية الائتمان تعبر عن الثقة التي تنشأ بين المقرض والمقترض والتي تدل على توافقهم في الرغبة والحاجات بين من تتوفر لديه الأموال ومن يحتاج إليها.

وبزيادة الائتمان يستدعي الى زيادة في حجم الموارد المتاحة للتوظيف والعكس فإن نمو حجم الموارد، يتطلب بحثاً عن فرص جديدة للائتمان تكون مضمونة وذات عائد مقبول، وهكذا يكون الادخار مقابلاً للائتمان له تأثير على سلوك كل من المدينين والدائنين سواء بواسطة جهاز الفائدة أم عن طريق السياسة الائتمانية والعوامل المؤثرة عليها(ذيب و آخرون، 2011، ص17)²⁰.

ويمكن النظر الى أهمية الائتمان المصرفي من عدة زوايا (طه، 2013، ص59)²¹:

1- أهمية الائتمان بالنسبة للبنك: الائتمان هو المصدر الأساسي الذي يستعمل في الاستثمار ومن خلاله

يستطيع البنك أن يحقق عوائد يضمن من خلالها الاستثمارية والنمو ويضمن القدرة على تحقيق مجموعة من الأهداف التي يسعى الى تحقيقها، ويتحقق ذلك نظراً لأن الائتمان بمفهومه العام يعتبر شكلاً من أشكال الاستثمار المصرفي بأصول البنك.

2- أهمية الائتمان بالنسبة للمقترضين (أفراد أو شركات): الحصول على القروض والتسهيلات

المصرفية يمكن المقترض من تغطية العجز المالي الذي قد يشل حركة نشاطه، فهو بذلك يفتح المجال أمام حركة الإنتاج والنمو في مجالات العمل المختلفة ويمكن الوحدات الاقتصادية من تحقيق أهدافها والاستمرارية في ممارسة أعمالها.

3- أهمية الائتمان بالنسبة للمجتمع: إن عملية منح الائتمان تساعد على خلق فرص استثمارية جديدة أو

التوسع في الأنشطة الحالية، وبالتالي يترتب على ذلك زيادة في الإنتاج والخدمات والتي تؤدي إلى فتح مناصب عمل جديدة وزيادة مستوى الدخل لأفراد المجتمع وتحقيق المزيد من الرفاهية الاقتصادية.

وبذلك اعتبر الائتمان المصرفي من أهم مصادر إشباع الحاجات التمويلية للقطاعات الاقتصادية المختلفة ولكن على الجانب الآخر يعتبر أداة حساسة قد تؤدي الى حدوث أضرار كبيرة بالاقتصاد إذا لم يحسن استخدامه.

²⁰ ذيب، سوزان سمير وزملاؤه. (2011) إدارة الائتمان. الطبعة الأولى، عمان، الأردن: دار الفكر، ص 17 .
²¹ طه، عمر هاشم. (2013) "دور سياسات منح الائتمان المصرفي في تقليل المخاطر وزيادة الأرباح(دراسة ميدانية في مصرف الشمال للتنمية والاستثمار)، مجلة جامعة الكركوك للعلوم الإدارية والاقتصادية، العدد 2، المجلد 3، ص 59.

3- أشكال الائتمان المصرفي:

يأخذ الائتمان المصرفي أشكالاً متعددة حيث يمكن تصنيفها في مجموعات مختلفة، وذلك وفقاً لما يلي (كراجة و آخرون، 2006)²²:

1- من حيث المدة: وله ثلاثة أشكال:

- الائتمان قصير الأجل: تتراوح مدته بين 30 يوم ولا تزيد عن عام، ويستخدم عادةً لسد النقص في رأس المال العامل ولمواجهة النفقات الجارية للمشاريع، يتميز عادة بانخفاض سعر الفائدة.
- الائتمان متوسط الأجل: تتراوح مدته بين سنة وخمسة سنوات على الأكثر، ويستخدم لأغراض التوسع الصناعي كسواء الآليات والمعدات أو تمويل الصادرات.
- الائتمان طويل الأجل: وهنا تزيد مدته عن خمس سنوات، ويستخدم في تمويل الاستثمارات في الأصول الثابتة، كسواء الإنشاءات والتجهيزات والأراضي. وبسبب طول أجله وزيادة المخاطر المحيطة بهذا الائتمان فإن كلفته المتمثلة بسعر الفائدة تكون مرتفعة.

2- من حيث الغرض منه: تمنح القروض المصرفية أو التسهيلات الائتمانية عادةً بأحد الأشكال التالية:

- الائتمان الاستثماري: وهو مخصص للإنتاج، حيث تمنح الأموال لتأسيس المشاريع أو توسيع المؤسسات الإنتاجية.
- الائتمان التجاري: وهو يهدف إلى تنشيط تداول السلع وتسهيل العمليات التجارية من خلال تمويل رأس المال العامل للأشخاص والشركات التجارية.
- الائتمان الاستهلاكي: وهي القروض الشخصية بغرض سد الحاجات الاستهلاكية مثل شراء السيارة أو الأثاث.

3- من حيث الضمانات المرتبطة به: ويكون الضمان إما شخصي أو عيني:

- فالائتمان الشخصي هو الائتمان الذي يحصل عليه المقترض مقابل وعد يقطعه على نفسه للمقرض دون أن يقدم أموال عينية للضمان، أو مقابل الوثوق بشخص يضمنه.
- أما الائتمان العيني فهو الائتمان الذي يحصل عليه المقترض مقابل تقديمه أموالاً عينية أو متداولة ضماناً لتسديد هذا القرض.

4- من حيث الشخص المقترض: يقسم إلى ائتمان مصرفي خاص وائتمان مصرفي عام، حيث أن:

- الائتمان المصرفي الخاص: وهو الذي يمنح لأشخاص القانون الخاص، كالأفراد الطبيعيين والأشخاص الاعتباريين كالشركات الخاصة.
- الائتمان المصرفي العام: وهو الذي يمنح لأشخاص القانون العام (الدولة والهيئات والمؤسسات العامة والمصالح الحكومية).

²² كراجة ، عبد الحليم .ربابعة ، علي .السكران، ياسر .مطر، موسى.(2006). الإدارة والتحليل المالي. الطبعة 1 ، عمان :دار صنعاء للنشر والتوزيع.

- من حيث طريقة استخدامه في البنوك: ويقسم الى ائتمان مباشر وائتمان غير مباشر (الزبيدي، 2001) 23:

أ- **الائتمان المصرفي المباشر:** وهو من الأكثر الأنواع شيوعاً في نشاط البنوك التجارية والأكثر ربحية ويتم هذا الائتمان بمنح مبالغ نقدية مباشرة لطالب الائتمان لاستخدامها في تمويل عمليات متفق عليها ومحددة مسبقاً بعقد الائتمان ومن أهم أشكال الائتمان النقدي المباشر:

- **الحساب الجاري المدين:** هو عبارة عن اتفاق بين البنك وطالب الائتمان بحيث يقدم البنك تسهيلات لعملائه وفي حدود سقف معين يستطيع هذا الأخير السحب من هذا الحساب وفي حدود السقف الممنوح يتم حساب الفائدة على المبلغ المستخدم.

- **القروض والسلفيات النقدية:** يعتبر من الأكثر الأنواع انتشاراً، يُلجأ إلى هذا النوع لتغطية احتياجات المقترضين المختلفة المستخدمة في تمويل احتياجات رأس المال العامل أو في عمليات التمويل الاستثماري طويل الأجل، قد تحمل الائتمانات نوع من الأنواع التالية:

1- القروض العادية: تتمثل في تلك القروض التي يمتد تاريخ استحقاقها إلى أكثر من سنة وتسدد على دفعة واحدة أو على عدة أقساط.

2- القروض المتجددة: يكون هذا القرض في شكل كميالية تستحق بعد بضعة شهور ولكن العميل له حق التجديد مرة أو أكثر.

3- قروض الحد الأعلى: هو عبارة على اتفاق يتم بين البنك والعميل بحيث يمنح له قرض له حد أقصى لفترة محددة على ألا يتجاوزه ويستطيع في المقابل ذلك أن يعيد الاقتراض ما دام قد سدد القرض السابق أو جزء منه.

- **الكمبيالات المخصومة:** الكمبيالات عبارة عن نوع من الأوراق التجارية الشائعة الاستعمال لدى البنوك التجارية بهدف تنظيم علاقات البيع الآجل فيما بين التجار، فهي عبارة على سند يتعهد بمقتضاها أحد التجار بدفع مبلغ معين إلى تاجر آخر في تاريخ محدد، لكن البنك يتدخل مقابل عمولة متفق عليها ليتمكن التاجر من خصم الكميالية قبل تاريخ استحقاقها.

ب- **الائتمان المصرفي غير المباشر:** يختلف هذا النوع من الائتمان عن الائتمان المباشر بأنه لا يعطي حقاً لطالب الائتمان باستعمال النقد بشكل مباشر. كما أنه لا يمثل ديناً مباشراً على العميل تجاه البنك.

والأنواع التالية هي الأكثر شيوعاً لصور الائتمان المصرفي غير المباشر:

- **الكفالات المصرفية (خطابات الضمان):** وهي عبارة عن عقد كتابي يتعهد بمقتضاه البنك (مصدر الكفالة) بضمان أحد عملائه بناء على طلبه في حدود مبلغ معين ولمدة معينة اتجاه طرف ثالث (المستفيد) بمثابة التزام ملقى على عائق العميل المكفول وضمناً لوفائه بالتزامه اتجاه المستفيد.

23 الزبيدي، محمود. (2001). إدارة الائتمان المصرفي والتحليل الائتماني. الطبعة 1، عمان: دار الوراق للنشر والتوزيع.

- **الاعتماد المستندي:** هو عقد يتعهد البنك بموجبه بدفع مبلغ محدد للمستفيد إذا ما قدم هذا الأخير في أجل محدد ووثائق محددة.
- **القبولات المصرفية:** وهي نوع من أنواع الائتمان المصرفي يقوم بموجبها طالب الائتمان بإصدار سحب زمني على أحد البنوك التي يتعامل معها إذ يتعهد البنك بدفع مبلغ السحب في تاريخ الاستحقاق.
- **بطاقات الائتمان:** تعد بطاقات الائتمان شكلاً متطوراً من أشكال الائتمان المصرفي الذي تقدمه البنوك التجارية لعملائها والذي انتشر كثيراً في السنوات الأخيرة، وتسمى أيضاً بطاقة الملاءة، وهي تعطي لحاملها الحق في التعامل مع العديد من المحلات التجارية المتفقة مع البنك المصدر للبطاقة على قبول منح الائتمان لتسديد مشترياته على أن يقوم بسداد قيمة المشتريات إلى البنك خلال فترة محددة من تاريخ استلامه لفاتورة الشراء مقابل عمولة معينة.
- **الائتمان الدولي:** تعمل البنوك على توسيع تعاملاتها الدولية لتسهيل التجارة الخارجية من خلال تمويل المعاملات الدولية سواء ذات الأجل الطويل أو القصير فالبنوك التجارية تساهم في تقديم الائتمان الدولي قصير الأجل من خلال ضمان الالتزامات الدولية للتجارية للزبائن وأوراق الخصم الدولية ... الخ.

4-العوامل المؤثرة في قرار منح الائتمان:

من أجل تقييم وإدارة المخاطر، تقوم إدارة الائتمان بدراسة فعالة لطلبات الائتمان المقدمة من قبل عملاء البنك قبل اتخاذ القرار الائتماني بالرفض أو بالقبول، في ضوء العناصر الائتمانية الحاكمة للنشاط الائتماني، حيث يتطلب كل قرار ائتماني اجراء موازنة بين العائد المتوقع وبين المخاطر والتكلفة المحتملة للائتمان المطلوب (Amdiu et al,2006,p93)²⁴.

لذا فإن اتخاذ القرار الائتماني يتطلب دراسة تحليلية لكل العوامل التي يمكن أن تؤثر على هيكل رأس المال (Capital Structure)، وقرارات الاستثمار (Investment decisions)، ومخاطر محفظة القروض (Loan portfolio risks). ويجب أن يقوم ضابط الائتمان بدراسة الملف الائتماني للعملاء من خلال جمع وتحليل المعلومات المتعلقة بالعميل، وبهذا يستطيع الموازنة بين العائد المتوقع وبين المخاطر والتكلفة

²⁴ Amidu, Mohammed and Hinson, Robert (2006). **Credit Risk, Capital Structure and Lending Decisions of Banks in Ghana** . Banks and Bank Systems, Vol. 1, No.1 , P.93

المحتملة للائتمان المطلوب، والوقوف على مدى مطابقة الحالة الائتمانية للسياسة الائتمانية للبنك من عدمه (Cebenoyan et al, 2004, p19-43)²⁵.

إن عملية تحليل المعلومات والبيانات عن حالة العميل المحتملة سوف تخلق القدرة لإدارة الائتمان في صنع القرار الائتماني، لذلك فإن إدارة الائتمان المصرفي تقوم بدراسة وتحليل عدد من المعايير الائتمانية التي من خلالها تستطيع الحكم واتخاذ قرار بمنح أو عدم منح الائتمان المصرفي للعميل.

ويمكن حصر العوامل المحددة لقرار الائتمان المصرفي فيما يلي (الدغيم وآخرون، 2006)²⁶:

1- العوامل المتعلقة بالعميل: ترتبط هذه العوامل بهدف واحد هو قياس مدى المخاطر التي قد يتعرض لها البنك نتيجة منحه لعميل معين من العملاء، وهل يمكن للبنك تحمل هذه المخاطر، ومدى سلامة الموقف الائتماني للعميل، لذلك فإن إدارة الائتمان المصرفي تقوم بدراسة وتحليل عدد من المعايير الائتمانية المتعلقة بالعميل كاستخدام نموذج معايير الائتمان المعروف 5C's:

والتي تعني أن هناك خمسة عناصر واجبة الدراسة جميعها تبدأ بالحرف C باللغة الإنجليزية والتي طبقاً لها يقوم المصرف كمانح ائتمان بدراسة تلك الجوانب لدى عميله المقترح كمقترض أو كعميل ائتمان. وفيما يلي استعراض لهذه العناصر (الخطيب، 2004)²⁷:

1- الشخصية (Character): تعد شخصية العميل الركيزة الأساسية الأولى في القرار الائتماني وهي الركيزة الأكثر تأثيراً في المخاطر التي تتعرض لها البنوك، وبالتالي فإن أهم مسعى عند إجراء التحليل الائتماني هو تحديد شخصية العميل بدقة.

فكلما كان العميل يتمتع بشخصية أمينة ونزيهة وسمعة طيبة في الأوساط المالية، وملتزماً بكافة تعهداته وحريصاً على الوفاء بالتزاماته، كلما كان أقدر على إقناع البنك بمنحه الائتمان المطلوب والحصول على دعم المصرف له.

يُعد قياس عامل الأمانة والنزاهة بدرجة دقيقة أمر تكتنفه بعض الصعوبات من الناحية العملية، ويتم التغلب على هذه الصعوبات من خلال الاستعلام الجيد وجمع البيانات والمعلومات عن العميل من المحيطين العملي والعائلي له، لمعرفة المستوى المعيشي وموارده المالية والمشاكل المالية التي يعانيها، ومستواه الاجتماعي وسجل أعماله التي قام بها وماضيه مع البنك ومع الغير وسابق تصرفاته مع البنوك الأخرى. ويتم ذلك عن طريق الاتصال بالشركة والعاملين بها، وبمورديها والبنوك التي سبق للعميل المقترح التعامل معها.

²⁵ Cebenoyan, A. Sinan and Strahan, Philip E. (2004) "Risk management, capital structure and Lending at banks", Journal of Banking & Finance, Vol. 28, P. 19-43.

²⁶ الدغيم، عبد العزيز و الأمين، ابراهيم وأنجرو، إيمان. (2006) "التحليل الائتماني ودوره في ترشيد عمليات الإقراض المصرفي"، مجلة جامعة تشرين للدراسات والبحوث العلمية، سلسلة العلوم الاقتصادية والقانونية المجلة، العدد الثالث، ص 28.

²⁷ الخطيب، منال. (2004) "تكلفة الائتمان المصرفي وقياس مخاطره بالتطبيق على أحد المصارف التجارية السورية". رسالة ماجستير، كلية الاقتصاد، جامعة حلب.

يوجد مجموعة من المؤشرات التي يتعين الحصول على بيانات كافية عنها أهمها ما يلي (الدغيم، 2006):²⁸:

- مدى انتظام العميل في سداد التزاماته مع البنوك الأخرى التي يتعامل معها ومن ثم يجب الحصول على استعلام عن العميل من هذه البنوك.
- مدى وفائه بتعهداته مع الشركات والمؤسسات والهيئات التي يتعامل معها سواء كان مورداً لاحتياجاته أو مستخدماً لمنتجاته.
- الاطلاع على سجلات المحاكم للوقوف على وجود أو عدم وجود حجوزات موقعة عليه أو أحكام لإخلاله بالتزاماته أو عدم الدفع.
- الوقوف على السجلات التجارية والأحكام التجارية والخاصة لمعرفة هل سبق إشهار إفلاس العميل أم لا.
- التصرفات الشخصية للعميل سواء في حياته الاجتماعية أو في محيط العمل، ومدى تمتعه باحترام المجتمع له.

فجوهر هذا الجانب في التحليل إذاً ينصرف إلى الاطمئنان إلى سمعة العميل على المستوى الشخصي ومستوى نشاطه ومدى رغبته واحترامه للوفاء بالتزاماته تجاه الآخرين.

2- القدرة (Capacity): وهي تعني إمكانية قيام العميل بسداد ما عليه من أقساط في المواعيد المحددة وذلك من خلال الوقوف على قدرة النشاط على توليد الدخل، ويمكن الاستدلال على ذلك من خلال مدى خبرة العميل في ممارسة نشاطه ويتم قياسها بعدد سنوات ممارسته لنشاطه بنجاح، ومقدار الأرباح المحققة من النشاط، وحجم التدفقات النقدية لنشاطه، وقابلية تحويل أصوله إلى نقدية.

ولابد للبنك عند دراسة هذا المعيار من استقراء العديد من المؤشرات التي تعكسها القوائم المالية الخاصة بالعميل وتحليلها، وهذه المؤشرات هي (David, 2003):²⁹:

- مؤشر نسبة التداول.
- مؤشر نسبة السيولة السريعة.
- مؤشر نسبة هامش الربح.
- حجم التدفقات النقدية.

3- رأس المال (Capital): يعتبر رأس مال العميل أحد أهم أسس منح القرار الائتماني، وعنصراً أساسياً من عناصر تقليل المخاطر الائتمانية باعتباره يمثل ملاءة العميل المقترض وقدرة حقوق

²⁸ الدغيم، عبد العزيز و الأمين، إبراهيم و أنجرو، إيمان. (2006) " التحليل الائتماني ودوره في ترشيد عمليات الإقراض المصرفي"، مجلة جامعة تشرين للدراسات والبحوث العلمية، سلسلة العلوم الاقتصادية والقانونية المجلة، العدد الثالث، ص 28.

²⁹ David E Vince. (2003) "Financial analysis and decision making". McGraw-Hill Companies. USA

ملكيتته على تغطية القرض الممنوح له، فهو بمثابة الضمان الإضافي في حال فشل العميل في التسديد.

ويرتبط هذا العنصر بمصادر التمويل الذاتية أو الداخلية للشركة والتي تشمل كل من رأس المال المستثمر والاحتياطات المكونة والأرباح المحتجزة، وجوهر هذا الجانب إذن ينصرف إلى أهمية التأكد من توافر رأس المال المناسب لدى العميل طالب الائتمان وقدرة حقوق ملكيته على تغطية الائتمان المطلوب، ومن ثم استعادة البنك ما سبق ومنحه للعميل في حال فشل العميل بذلك.

4- الضمان (Collateral): يقصد بالضمان مجموعة الأصول التي يضعها العميل تحت تصرف البنك كضمان مقابل الحصول على القرض، ولا يجوز للعميل التصرف في الأصل المرهون، فهذا الأصل سيصبح من حق البنك في حال عدم قدرة العميل على السداد.

وإذا ما كان القرار الائتماني قد استلزم تقديم ضمانات معينة، فإنه يجب أن يراعى عند تحديد الضمان ما يلي (الدغيم وآخرون، 2006)³⁰:

- أن يكون قابلاً للبيع أو التصفية ولا يمثل الاحتفاظ به تكلفة أو عبئاً مرتفعاً.
- عدم تقلب قيمة الضمان بشكل كبير خلال فترة الائتمان، فالضمان الذي يتعرض لتقلبات شديدة في قيمته لا يمكن الاعتماد عليه في استرداد حقوق البنك.
- أن تكون ملكية العميل للضمان ملكية كاملة وليست محل نزاع.
- أن يتوفر لدى البنك كافة المستندات القانونية التي تؤكد حق البنك باستخدام هذه الضمانات وتسييلها وذلك لسداد قيمة الائتمان عند عجز العميل عن السداد.
- كفاية الضمانات المقدمة لتغطية قيمة الائتمان والفوائد والعمولات الأخرى قدر المستطاع.

5- الظروف المحيطة (Conditions): ويقصد بها مجموعة الظروف العامة والخاصة التي يتأثر بها نشاط العميل أو الشركة.

فالظروف العامة ترتبط بالتضخم الاقتصادي، ودورة الأعمال والأنشطة الاقتصادية، وتغيرات سعر صرف العملات الأجنبية، وكذلك الإطار التشريعي والقانوني الذي تعمل الشركة في إطاره الخارجي، حيث تؤثر هذه الظروف العامة في مختلف قطاعات النشاط الاقتصادي.

³⁰ الدغيم، عبد العزيز و الأمين، ابراهيم و أنجرو، إيمان. (2006) " التحليل الائتماني ودوره في ترشيد عمليات الإقراض المصرفي"، مجلة جامعة تشرين للدراسات والبحوث العلمية، سلسلة العلوم الاقتصادية والقانونية المجلة، العدد الثالث، ص 28.

أما الظروف الخاصة ترتبط بالنشاط الخاص الذي يمارسه العميل، مثل الحصة السوقية لمنتجات العميل أو الخدمات التي يقدمها، وموقع المشروع من دورة حياته بمعنى هل هو في مرحلة التقديم أو الولادة، أم في مرحلة النمو، أم في مرحلة الاستقرار، أم في مرحلة الانحدار.

ومما يرتبط بالظروف الخاصة أيضاً، نوع السلعة أو الخدمة التي يقدمها العميل ومواصفاتها ومدى جودتها وقابليتها للتلف، ومتوسط تكلفة السلعة أو الخدمة وسعرها السائد في السوق. وكذلك حجم الطلب عليها. وكذلك حجم السوق فيما يتصل بعدد المستهلكين وخصائصهم فيما يتعلق بالسن، والدخل، وسلوكهم الاستهلاكي. وكذلك منافذ التوزيع من حيث الحجم، والنوع، والعدد، والانتشار الجغرافي (الدغيم وآخرون، 2006)³¹.

2- العوامل المتعلقة بالبنك: ويتسع هذا المجال ليشمل العوامل الآتية:

- درجة السيولة التي يتمتع بها البنك حالياً وقدرته على توظيفها، ومفهوم السيولة يعني قدرة البنك على مواجهة التزاماته، والتي تتمثل بصفة أساسية في عنصرين هما: تلبية طلبات المودعين للسحب من الودائع، وأيضاً تلبية طلبات الائتمان، أي القروض والسلفيات لتلبية احتياجات المجتمع.
- نوع الاستراتيجية التي يتبناها البنك في اتخاذ قراراته الائتمانية ويعمل في إطارها، أي في استعداده لمنح ائتمان معين أو عدم منح هذا الائتمان.
- الهدف العام الذي يسعى البنك إلى تحقيقه خلال المرحلة القادمة.
- القدرات التي يمتلكها البنك وخاصة الكوادر البشرية المؤهلة والمدرّبة على القيام بوظيفة الائتمان المصرفي، وأيضاً التكنولوجيا المطبقة وما يمتلكه البنك من تجهيزات إلكترونية حديثة.

3- العوامل الخاصة بالائتمان المطلوب: ويمكن حصر هذه العوامل بما يلي:

- المدة الزمنية التي يستغرقها الائتمان، أي المدة التي يرغب العميل بالحصول على الائتمان خلالها، ومتى سيقوم بالسداد وهل تتناسب فعلاً مع إمكانيات العميل.
- مصدر السداد الذي سيقوم العميل المقترض بسداد المبلغ منه.
- طريقة السداد المتبعة، أي هل سيتم سداد الائتمان دفعة واحدة في نهاية المدة، أم سوف يتم سداده على أقساط دورية، وذلك بما يتناسب مع طبيعة نشاط العميل ومع إيراداته وموارده الذاتية وتدفقاته الداخلة.
- الغرض من الائتمان ونوعه وهل يتوافق مع السياسة العامة للائتمان في البنك أم يتعارض معها.

³¹ دغيم، عبد العزيز و الأمين، إبراهيم و أنجرو، إيمان. (2006) "التحليل الائتماني ودوره في ترشيد عمليات الإقراض المصرفي"، مجلة جامعة تشرين للدراسات والبحوث العلمية، سلسلة العلوم الاقتصادية والقانونية، العدد الثالث، ص 28.

- قيمة هذا الائتمان، وله أهمية خاصة، حيث إنه كلما زاد المبلغ عن حد معين كان البنك أحرص في الدراسات التي يجريها خاصةً أن نتائج عدم سداد قرض بمبلغ ضخم تكون صعبة وقد تؤثر على سلامة المركز المالي للبنك.

5-مخاطر الائتمان المصرفي:

1.5. مفهوم المخاطر الائتمانية:

تعتبر مخاطر التخلف عن سداد القرض واحدة من أهم مخاطر الائتمان في قرارات الاستثمار وتتميز قرارات الاستثمار عموماً سواء كانت في استثمارات عينية او استثمارات مالية بعدم قدرة المستثمر على تحديد العائد المتوقع لذلك الاستثمار على وجه اليقين. ويرجع ذلك إلى حالة عدم القدرة على الإحاطة في شكل المستقبل الأمر الذي يجعل الاستثمار محاطاً بدرجة من المخاطر. وتنشأ مخاطر الائتمان عندما لا يستطيع البنك تقييم قدرة العميل على الوفاء بالتزاماتها بتسديد أصل القرض وفوائده.

وتعرف المخاطر الائتمانية بأنها هي الخسائر التي يمكن أن يتحملها البنك بسبب عدم قدرة الزبون أو عدم وجود النية لديه لسداد أصل مبلغ الائتمان(رضا وآخرون،1999)³².

وعرفت أيضاً بأنها المخاطر الحالية والمستقبلية التي يمكن أن تتأثر بها إيرادات البنك ومركزه المالي، والناجمة عن عدم قيام العميل بالوفاء بالتزاماته تجاه البنك بالوقت المناسب(كراسنة،2006)³³.

2.5. أنواع المخاطر الائتمانية:

تتنوع مخاطر الائتمان وفقاً لتنوع مصادرها، وبناء عليه يمكن حصر أنواعها وفقاً لما يلي(الزبيدي،2001)³⁴:

- 1- مخاطر العميل أي تلك المتعلقة بطبيعة نشاط الشركة ونوعية إدارتها وأدائها التشغيلي.
- 2- مخاطر الصناعة وهي على علاقة بنوع الصناعة التي تعمل فيها الشركة وبنوع منتجاتها.
- 3- مخاطر السيولة والتي ترتبط باحتمال تعثر العميل بعد منح الائتمان، مما يؤدي الى عجزه عن تسديد أصل القرض وفوائده.

³² رضا، عبد المعطي. جودة، محفوظ أحمد.(1999) "إدارة الائتمان". عمان: دار وائل للنشر والتوزيع.
³³ كراسنة، إبراهيم.(2006). أطر أساسية و معاصرة في الرقابة على البنوك و إدارة المخاطر. أبوظبي:معهد السياسات الاقتصادية.
³⁴ الزبيدي، محمود. (2001). إدارة الائتمان المصرفي والتحليل الائتماني. الطبعة 1 ، عمان: دار الوراق للنشر والتوزيع.

- 4- مخاطر القطاع الاقتصادي الذي ينتمي إليه العميل: ترتبط هذه المخاطر بطبيعة النشاط الذي يعمل فيه العميل، إذ أن لكل قطاع اقتصادي مخاطر تختلف باختلاف الظروف التشغيلية والإنتاجية والتنافسية لوحدات هذا القطاع.
- 5- مخاطر الظروف العامة: ترتبط هذه المخاطر بالظروف الاقتصادية والتطورات السياسية والاجتماعية.
- 6- مخاطر مرتبطة بالبنك: ترتبط هذه المخاطر بمدى كفاءة إدارة الائتمان في البنك في متابعة الائتمان المقدم للعميل، والتحقق من قيام العميل بالشروط المتفق عليها في اتفاقية منح الائتمان. ومن هذه الأخطاء هي عدم قيام البنك بحجز ودائع العميل، والتي وضعها كضمان للتسهيلات الائتمانية وسحب العميل لهذه الودائع.

3.5. إدارة مخاطر الائتمان المصرفي:

يأتي دور إدارة مخاطر الائتمان في إبقاء التعرض للمخاطر الائتمانية ضمن معايير ومستويات آمنة، لتعظيم معدلات العوائد مقابل المخاطر المقبولة. واهتمت لجنة بازل كنظام رقابي للمخاطر المصرفية التي تواجه النشاط المصرفي وأنظمت قياس المخاطر، وطرق إدارتها للتخفيف منها والمحافظة على استقرار الجهاز المصرفي.

وتعتبر وظيفة إدارة المخاطر على مستوى البنك هي وظيفة أساسية تهتم بالتنسيق بين كافة الإدارات البنكية من حيث تحديد مختلف المخاطر التي تواجه البنك بشكل دوري وتبويبها في صورة تقارير ترفع الى الإدارة العليا لمناقشتها(بالعجوز،2005)³⁵.

وفي هذا السياق أصدرت لجنة بازل وثيقة تتعلق بمعايير إدارة مخاطر الائتمان حتى تشجع المراقبين المصرفيين على المستوى الدولي على تعزيز الممارسات السليمة لإدارة مخاطر الائتمان، وتتضمن هذه المعايير (Basel Committee,2000,p5)³⁶:

- 1- **توافر مناخ ملائم لإدارة المخاطر الائتمانية:** وتعني موافقة مجلس الإدارة على الاستراتيجيات الهامة لمنح الائتمان بالبنك وتتضمن الاستراتيجية لمنح الائتمان في:
- مدى استعداد البنك لتحمل المخاطر ومستوى الربحية نتيجة ذلك.

³⁵ بالعجوز، حسين.(2005).إدارة المخاطر البنكية والتحكم فيها.مداخلة مقدمة الى الملتقى الوطني حول المنظومة المصرفية في الالفية الثالثة،جامعة جيجل الجزائر.ص.9.

³⁶ Basel Committee. (2000) "Principles for the Management of Credit Risk"

- تحديد أنواع الائتمان المصرفي الممكن منحه للعملاء وتحديد القطاعات والمناطق الجغرافية الممكن منحها للائتمان.
- وضع حدود قصوى لأجال منح الائتمان وأسس تسعير الائتمان.
- تحديد إرشادات عن (نسبة القروض الى الأصول أو نسبة كل نوع من أنواع الائتمان الى الأصول) و (نسبة كل نوع من أنواع الائتمان الى إجمالي المحفظة الائتمانية أو حقوق الملكية).
- تحديد نوعية الضمانات وكيفية تقديمها والجهة التي تقيمها والعلاقة بين حجم الائتمان وقيمة الضمانات.

2- توافر إجراءات سليمة لمنح الائتمان: ويمكن الإشارة في هذا المجال إلى المعايير الملائمة لمنح الائتمان وكذلك الحدود التي يتم الالتزام بها لدى منح الائتمان، ومن أهم المعايير الملائمة لمنح الائتمان هي مايلي:

- معلومات كافية لإجراء تقييم شامل لنوعية المخاطر المرتبطة بطالب القرض وفقاً لنظام التصنيف الداخلي بالبنك.
- الأهلية القانونية لطالب القرض لتحمل الالتزام وكذلك السمعة والخبرة والغرض من القرض.
- مصادر السداد والتدفقات النقدية المتوقعة.
- طبيعة المخاطر الحالية والمستقبلية كطالب القرض للصناعة، ومدى حساسية للتطورات الاقتصادية.
- مدى الالتزام بسداد الالتزامات السابقة.
- العلاقة بين المخاطر والربحية.

3- توافر إجراءات للتعامل مع الائتمان ومتابعته: حيث يجب العمل على توفير مايلي:

- توافر نظام للتعامل مع ملفات الائتمان ومستندات المديونية.
- متابعة تنفيذ الائتمان وفقاً لشروط الائتمان.
- التصنيف الداخلي للائتمان والذي يساعد على (منح الائتمان ومتابعة مدى جودة تسعير الائتمان-تحديد الخصائص المحفظة الائتمانية والتركزات – تحديد القروض المتعثرة ومدى كفاية المخصصات).

4- توافر إجراءات كافية للرقابة على مخاطر الائتمان: تتضمن الإجراءات الكافية للرقابة على مخاطر الائتمان في وجود مايلي:

- نظام مستقل لمراجعة الائتمان بهدف التعرف على مدى كفاءة المسؤولين عن منح الائتمان ومتابعته ومدى سلامة إجراءات التعامل مع الائتمان ومدى جودة المحفظة الائتمانية ومدى سلامة نظام التصنيف.

- الرقابة الداخلية للتأكد من الإبلاغ عن الاستثناءات في السياسات الائتمانية وفي الإجراءات الائتمانية وفي الحدود الائتمانية.
- وحدة خاصة لاكتشاف التسهيلات الائتمانية المتعثرة في وقت مبكر.

4.5. نظام تصنيف مخاطر الائتمان:

لعل من أهم الأدوات الفاعلة والكفوة التي تساعد دائرة الائتمان المصرفي في البنك أولاً في اتخاذ قرار منح الائتمان، وتحديد قيمة القرض وأجله، وسعر الفائدة، ثم بعد ذلك في متابعة مقدرة العميل على تسديد أصل القرض وفوائده، هو وجود نظام معتمد لتصنيف مخاطر الائتمان (الشيخلي، 2012)³⁷.

ومن أهم أهداف وجود نظام تصنيف مخاطر الائتمان هو التنبؤ فيما إذا كان طالب القرض قادر على السداد في الوقت المحدد أم لا، مما يساعد ضابط الائتمان في اتخاذ قرار المنح من عدمه (Banasik et al, 2003)³⁸.

وقد تم تحديد أسس لإعداد نظام التصنيف الائتماني الداخلي، وذلك من خلال تعليمات تم إصدارها من قبل مجلس النقد والتسليف-مصرف سورية المركزي والتي جاءت فيما يلي³⁹:

1- يتوجب على المصارف أن تصمم وتضع موضع التنفيذ نظاماً داخلياً لتصنيف مخاطر الائتمان، بحيث يشمل هذا النظام، كافة الطرق والأساليب والمعايير والعمليات والضوابط والبيانات المجمعة والبرمجيات التي تستخدم في إطار مراقبة جودة التعرضات الائتمانية وتقييم مخاطرها وتحديد مستوى هذه المخاطر وتقدير احتمال التعثر والخسائر.

2- يُراعى لدى تصميم نظام التصنيف الداخلي للتعرضات الائتمانية تجاه الحكومات والبنوك والشركات (Corporate) أن يُبنى على بعدين منفصلين (Two Separate and Distinct Dimensions) وفق الآتي:

- مخاطر تعثر العميل (المقترض/المصدر (Borrower/Issuer)) بناءً على الجدارة الائتمانية (Creditworthiness) واحتمال التعثر (Probability of Default) ، حيث يتم تصنيف كافة التعرضات الائتمانية (بغض النظر عن طبيعتها أو الخصائص المرتبطة بنوع التعرض الائتماني) تجاه ذات العميل ضمن ذات الدرجة (Same Grade) ، وعلى أن تتضمن سياسات وإجراءات

³⁷ الشيخلي، هديل. (2012) "العوامل الرئيسية المحددة لقرار الائتمان المصرفي في البنوك التجارية الأردنية". رسالة ماجستير، كلية الأعمال، جامعة الشرق الأوسط.

³⁸ Banasik J., Crook J., and Thomas L. (2003) "Sample selection bias in credit scoring models" Journal of the Operational Research Society, Vol 54, P. 822-832.

³⁹ تعميم رقم 4/م لعام 2019 الخاص بتطبيق بإلزام المصارف العاملة في سورية على تطبيق معيار التقارير المالية الدولي رقم 9.

التصنيف الائتماني الداخلي توضيحاً للمخاطر الائتمانية التي تعكسها كل درجة من درجات التصنيف لجهة المعايير والمؤشرات المستخدمة لتحديد مستوى هذه المخاطر واحتمال التعثر (PD) للعملاء المصنفين ضمنها، وبحيث يكون التغير في مستوى المخاطر ملموس وقابل للقياس مقابل تغيير الجودة الائتمانية بين كل درجة من درجات التصنيف الائتماني والدرجة التي تليها أو تسبقها.

- العوامل الخاصة بالتعرض الائتماني (Transaction-Specific Factors) مثل الضمانات وأولوية الدين (Seniority) ونوع المنتج الائتماني (Product Type) والغرض من التسهيل وأي عوامل أخرى يمكن أن تؤثر في احتساب الخسارة عند التعثر ((Loss Given Default (LGD)).

3- يُصمم نظام التصنيف الداخلي لتعرضات التجزئة الائتمانية بحيث يكون شاملاً لمخاطر تعثر العميل والعوامل الخاصة بالتعرض الائتماني دون أن يتم الفصل ما بين هذين البعدين، وعلى أن يتم تجميع التعرضات الميَّوبة ضمن فئات التجزئة في مجموعات (Pools/Groups)، بحيث تتكون كل مجموعة (Pool) من تعرضات ائتمانية متجانسة (Homogenous) وذات مخاطر متشابهة ومشاركة (Shared Credit Risk Characteristics)، بما يسمح بتقدير دقيق ومتسق لاحتمال التعثر (PD) والخسارة عند التعثر (LGD) أو معدّل التعثر/الخسارة التاريخي لكل مجموعة، وعلى أن يتم الأخذ بالاعتبار المعايير التالية كحد أدنى لدى تجميع التعرضات الائتمانية في كل مجموعة:

- خصائص المخاطر المرتبطة بالعميل، مثل نوع العميل وعمره ومكان إقامته.

- خصائص المخاطر المرتبطة بالتعرض الائتماني، مثل أولوية الدين ونوع المنتج والضمانات المرتبطة به ونسبة القرض إلى القيمة (LTV) وعمر التسهيل.

- التأخر عن السداد (Delinquency)، حيث يجب فصل الديون التي استحق أحد أقساطها (أصل الدين أو فوائد/عوائد) ولم يتم سداده في حينه في مجموعات مختلفة.

- عدم تجميع التعرضات الائتمانية بالطريقة التي تحول دون تمييز ارتفاع المخاطر الائتمانية لأحد التعرضات المجمعة في ظل أداء جيد للمجموعة ككل.

4- يجب أن تكون عملية التصنيف الائتماني شاملة لكافة التعرضات الائتمانية، وعلى المصارف استخدام أنظمة ومنهجيات/أساليب التصنيف (Rating Methodologies/ Systems) المناسبة ضمن كل فئة أو كل فئة فرعية من فئات التعرضات الائتمانية (Asset Class or Sub-class)، وأي فئات فرعية أخرى يرى المصرف ضرورة تعريفها وتحديدها، وذلك مع إمكانية تعدد المنهجيات ضمن الفئة الواحدة أو الفئة الفرعية الواحدة تبعاً لخصائص المخاطر التي من الممكن أن تختلف بحسب القطاع الاقتصادي أو بالنسبة لصناعات محددة ضمن القطاع الاقتصادي الواحد أو بحسب العملة أو أجل التعرض أو أي عوامل أخرى يرى المصرف ضرورة تمييزها، إلا أنه - وفي حال تعدد منهجيات التصنيف وفق المبيّن أنفاً - يتوجب توثيق الآلية التي يتم من خلالها توزيع

التعرضات الائتمانية على المنهجيات المستخدمة وتطبيق هذه الآلية بالطريقة التي تعكس بدقة مستوى المخاطر الائتمانية المرتبطة بكل تعرض.

5- في حال استخدام أنظمة تصنيف ائتماني داخلي على الأجل القصير (Short-Term Rating Systems) يتوجب موازنة هذه الأنظمة مع نظام التصنيف طويل الأجل بما يعكس تغير الجودة الائتمانية ومستوى المخاطر للتعرضات الائتمانية قصيرة الأجل بشكل دقيق.

6- يتكون نظام التصنيف الائتماني المبني على أساس مخاطر تعثر العميل من سبع درجات للتعرضات الائتمانية غير المتعثرة (Non-Defaulted) وثلاث درجات للتعرضات الائتمانية المتعثرة (Defaulted)، بحيث يتم تعريف كل درجة على أنها تقييم محدد لمخاطر العميل بناءً على معايير واضحة ومعتمدة وموثقة يتم على أساسها تحديد مجال احتمال التعثر (PD band)، في حين يعود لكل مصرف – بناءً على أدلة تجريبية موثقة - تحديد عدد درجات نظام التصنيف الائتماني المبني على أساس العوامل الخاصة بالتعرض الائتماني، وعلى أن يكون عدد درجات هذا النظام كافٍ لتجنب تركيز تعرضات ذات مجال واسع في مستوى الخسارة عند التعثر (LGD) في درجة واحدة.

7- دون الإخلال بدور دائرة إدارة المخاطر، يمكن أن يكون ضباط الائتمان المسؤولون عن مسك الملفات الائتمانية (Front-Line Lending Staff) مسؤولون أيضاً عن التحديد الأولي لدرجة التصنيف الائتماني لكل عميل (Corporate) أو مجموعة من عملاء التجزئة (Pool) عند الاعتراف الأولي، إلا أن عملية التصنيف المنفذة من قبلهم يجب أن تتم وفقاً لأدلة إجراءات تفصيلية معتمدة وموثقة تتضمن ضوابط خاصة بإصدار الآراء الحكمية (Judgment) وأن تخضع هذه العملية بشكل لاحق للمراجعة والتدقيق من قبل وظيفة مستقلة.

8- يتم تحديث التصنيف الائتماني لكافة العملاء والتعرضات الائتمانية ومجموعات تعرضات التجزئة وكذلك مدى اتساق خصائص تعرضات التجزئة المضمنة ضمن المجموعة الواحدة على أساس سنوي على الأقل، وذلك دون الإخلال بوجود تعديل التصنيف/التجميع عندما تتاح أية معلومات جديدة ذات صلة أو لها تأثير على احتساب معاملات الخسائر الائتمانية المتوقعة وفق متابعات المصرف الدورية والمستمرة في هذا الإطار.

9- يُراعى ألا ينتج عن توزيع (العملاء/التعرضات الائتمانية)؛ على درجات نظام التصنيف الائتماني تركزاً هاماً في أي درجة من هذه الدرجات، وفي حال حدوث مثل هذا التركيز يتوجب مراجعة سياسات وإجراءات التصنيف والتأكد من أنّ مجال احتمال التعثر أو الخسارة عند التعثر (PD/LGD band) المرتبط بهذه الدرجة يغطي كافة التعرضات المصنفة ضمنها.

10- لدى استخدام المصرف للنماذج الإحصائية أو أية نماذج أخرى لتحديد درجات المخاطر الائتمانية (Credit Scoring) لغايات التصنيف أو احتساب معاملات الخسائر الائتمانية المتوقعة (PD, LGD, EAD)، يتوجب توثيق آليات وطرق استخدام هذه النماذج بما في ذلك كافة الفرضيات

والطرق الرياضية وكذلك عمليات اختبار هذه النماذج وفعاليتها وتقدير الأخطاء التي قد تنجم عن استخدامها، وفي هذه الحال لا يجوز تجاهل أو تعديل مخرجات هذه النماذج بموجب تقييم أو خبرة إلا في حالات محددة يمكن أن تنص عليها أدلة سياسات وإجراءات التصنيف وعلى أن تخضع لموافقة مجلس الإدارة أو أحد اللجان المنبثقة عنه.

11- يتوجب على المصارف الاحتفاظ بسجلات التصنيف الائتماني التاريخية وكافة المعلومات المستخدمة في عملية التصنيف واحتساب معاملات الخسائر الائتمانية المتوقعة سواء للحكومات أو البنوك أو الشركات أو مجموعات التعرضات الائتمانية للتجزئة، وذلك بشكل دائم ومستمر لأغراض رصد التاريخ الائتماني للعملاء و/أو لمجموعات التعرضات الائتمانية (Pools) وكذلك لأغراض المراجعة والتدقيق ورفع التقارير التي قد تطلبها مديرية مفوضية الحكومة لدى المصارف.

ثانياً: التنقيب في البيانات:

1- مفهوم وأهمية التنقيب في البيانات:

ظهر مصطلح التنقيب في البيانات في منتصف التسعينات في الولايات المتحدة الأمريكية، وهو يجمع ما بين الإحصاء وتكنولوجيا الإعلام (قواعد البيانات، الذكاء الاصطناعي، التعلم الآلي)، وتوجد عدة تعريفات لهذا المفهوم منها، حيث يمكن تعريفها بأنها: " الاستكشاف الآلي أو المؤتمت لأنماط شائعة و غير جلية مخفية في قاعدة بيانات معينة" (العلاق، 2005، 40)، أو أنها: " سيرورة تحليل دقيقة وذكية، تفاعلية و تسلسلية، تسمح لمديري النشاطات عند استخدام هذه السيرورة باتخاذ قرارات والقيام بأعمال ملائمة في صالح النشاط المسؤولين عنه و المؤسسة التي يعملون بها" (Bazszlica, 2001، 41)، أو أنها: " عبارة عن تحليلات لكمية كبيرة من البيانات بغرض إيجاد قواعد و أمثلة و نماذج التي يمكن أن تستخدم تقود و تدل أصحاب القرار، و تتنبأ بالسلوك المستقبلي" (العلي وآخرون، 2006، 42)، كما يمكن تعريفها كذلك بأنها: " تحليل لمجموعات كبيرة الحجم من البيانات المشاهدة للبحث عن علاقات محتملة و تلخيص للبيانات في أشكال جديدة لتكون مفهومة و مفيدة لمستخدمها" (Hand et al, 2001، 43). وكما عرفت أيضاً على أنها: " نشاط يستخرج بعض المعلومات الجديدة الهامة الموجودة في قواعد البيانات أو في مستودع البيانات" (العلاق، 2005، 44).

40 العلاق، بشير عباس . الإدارة الرقمية: المجالات و التطبيقات، مركز الإمارات للدراسات و البحوث الإستراتيجية، ابوظبي، 2005، ص 84.

41 Bazsalica M., Naim P., Data mining pour le Web, éd. Eyrolles, Paris, 2001, P. 61.

42 العلي، عبد الستار. قنديلجي، عامر إبراهيم. العمري، غسان. المدخل إلى إدارة المعرفة، دار المسيرة للنشر و التوزيع و الطباعة، الطبعة الأولى، عمان، 2006، ص. 157.

43 Hand d., Mannila H., Smyth R., Principles of Data Mining, MIT Press, London, 2001, p. 01.

44 العلاق، بشير عباس. مرجع سابق، ص. 89-91.

تم تحديد عملية اكتشاف المعرفة الشاملة كعملية تفاعلية ومتكررة تتألف من الخطوات التالية: فهم مجال التطبيق، واختيار البيانات، وتنظيف البيانات والمعالجة المسبقة، وتكامل البيانات، واختيار خوارزميات استخراج البيانات، واستخراج البيانات وتفسير النتائج واستخدام المعرفة المكتشفة (Fayyad et al, 1996)⁴⁵.

ويقوم التنقيب عن البيانات بأتمتة اكتشاف الأنماط ذات الصلة في قاعدة البيانات، باستخدام مناهج وخوارزميات محددة للنظر في البيانات الحالية والتاريخية التي يمكن تحليلها بعد ذلك للتنبؤ بالاتجاهات المستقبلية. ونظراً لأن أدوات التنقيب عن البيانات تتنبأ بالاتجاهات والسلوكيات المستقبلية من خلال قراءة قواعد البيانات بحثاً عن الأنماط المخفية، فإنها تسمح للمؤسسات باتخاذ قرارات استباقية قائمة على المعرفة (Silltow, 2006)⁴⁶.

كما تأتي أهمية التنقيب في البيانات حيث أصبحت تكنولوجيا مفيدة ومهمة في ظل التطور السريع وانتشار واستخدام قواعد البيانات كما أن استخدامها يوفر للمؤسسات في جميع المجالات القدرة على الاستكشاف والتركيز على أهم المعلومات في قواعد البيانات. تركز تقنيات التنقيب على بناء تنبؤات مستقبلية واستكشاف السلوك والاتجاهات، مما يسمح بتقدير قرارات صحيحة واتخاذها في الوقت المناسب، غير أنها تجيب على العديد من الأسئلة في وقت قياسي خاصة تلك النوعية من الأسئلة التي يصعب الإجابة عليها، إن لم يكن مستحيلاً، باستخدام تقنيات الإحصاء الكلاسيكية (سليمان، 2017)⁴⁷.

2- أهداف التنقيب في البيانات:

هناك ثلاثة أهداف للتنقيب في البيانات (علي و آخرون، 2011، ص209-236)⁴⁸ :

- 1- من أجل تحليل بعض الظواهر المرئية.
- 2- من أجل التثبت من نظرية ما.
- 3- من أجل تحليل البيانات للحصول على علاقات جديدة وغير متوقعة.

⁴⁵ Fayyad, U. Piatetsky-Shapiro, G. & Smyth, P. (1996), —From data mining to knowledge Discovery in database||, American Association for Artificial Intelligence Press, Cambridge.

⁴⁶ Silltow J., (2006).Data mining tools and techniques. United Kingdom

⁴⁷ سليمان، الجبلي كباشي ابراهيم. (2017): " تنقيب بيانات المحلات التجارية بواسطة قاعدة الارتباط ". رسالة ماجستير غير منشورة، كلية الدراسات العليا، جامعة النيلين.

⁴⁸ علي، بسام، وزملاؤه (2011) . " استخدام تنقيب البيانات في التنبؤ بظاهرة تسرب طلبة الكلية بالتطبيق علي كلية الإدارة والاقتصاد"، المجلة العرفية للعلوم الاحصائية عدد(18)، ص209-236.

3-مراحل عملية التنقيب في البيانات:

يمكن تلخيص مراحل وخطوات عملية التنقيب في البيانات كما يلي(العلي و آخرون،2006)⁴⁹:

1- **فهم طبيعة الأعمال (Business Understanding)**: يعتبر المطلب الأول لاكتشاف المعرفة هو فهم المشاكل والمسائل التي تواجهها الأعمال. وبمعنى آخر، كيف يمكن تحقيق المنفعة العظمى من التنقيب في البيانات، مما يتطلب وجود صيغة واضحة ومحددة الأهداف.

2- **فهم البيانات (Data Understanding)** : تعتبر مسألة معرفة ماهية وطبيعة البيانات عامل مهم في نجاح عملية التنقيب في البيانات و اكتشاف المعرفة. حيث أن معرفة البيانات بصورة جيدة تعني مساعدة المصممين على استخدام الخوارزميات أو الأدوات المستخدمة للمسائل المحددة بدقة عالية. وهذا يقود إلى تعظيم فرص النجاح بالإضافة إلى رفع الفعالية والكفاءة لنظام اكتشاف المعرفة.

ولا تحتاج عملية التنقيب في البيانات إلى جميع البيانات في مستودع البيانات، أما إذا كان مستودع البيانات موجود في المنظمة، فمن الأفضل عدم احتكار المستودع بشكل مباشر لغرض التنقيب في البيانات.

ويمكن تلخيص الخطوات الضرورية لعملية فهم البيانات كالآتي:

- **تجميع البيانات (Data Collection)**: و هي الخطوة الموجهة نحو تحديد مصدر البيانات في الدراسة بما في ذلك استخدام البيانات العامة الخارجية مثل الضرائب و غيرها.
- **توصيف البيانات (Data Description)**: و هي الخطوة التي تركز على توصيف محتويات الملف الواحد من الملفات أو الجداول.
- **جودة البيانات وتحققها (Data Quality and Verification)**: هذه الخطوة تحدد ما إذا كان تقليل أو إهمال بعض البيانات غير الضرورية أو كونها رديئة الجودة و قد لا تنفع في الدراسة. لأن النموذج الجيد يحتاج إلى بيانات جيدة مما يتوجب أن تكون البيانات صحيحة وذات مضمون دقيق.
- **التحليل الاسترشادي للبيانات (Exploratory Analysis of Data)**: تستخدم الأساليب مثل الإظهار المرئي أو التصور أو عملية التحليل المباشر(OLAP) التي تؤدي إلى إجراء التحليل الأولي للبيانات. وتعتبر هذه الخطوة مهمة وضرورية لأنها تركز على تطوير الفرضيات المتعلقة بالمشكلة قيد الدراسة.

49 العلي،عبد الستار. قنديلجي، عامر إبراهيم. العمري، غسان. المدخل إلى إدارة المعرفة، دار المسيرة للنشر و التوزيع و الطباعة، الطبعة الأولى، عمان، 2006، ص. 157.

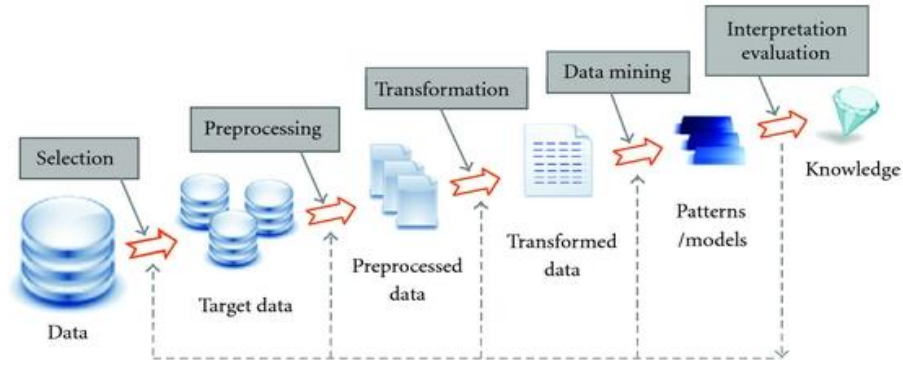
3- تهيئة البيانات (Data Preparation): و تشمل الخطوات التالية:

- الاختيار (Selection) وتعني اختيار المتغيرات المتوقعة و حجم العينة.
- صياغة المتغيرات وتحويلها (Construction and Transformation Variables) حيث يجب دائماً أن تصاغ المتغيرات الجديدة لبناء النماذج الفعالة.
- تكامل البيانات (Data Integration): حيث أن مجاميع البيانات في دراسة التنقيب عن البيانات من الممكن تخزينها في قواعد بيانات متعددة الأغراض التي تكون بحاجة إلى توحيدها في قاعدة بيانات واحدة.
- تصميم و تنسيق البيانات (Data Formatting) حيث تتعلق هذه الخطوة في إعادة ترتيب حقول البيانات كما يتطلب في نموذج التنقيب في البيانات.

4- صياغة نماذج الحل وثبوتها (Model Building and Validation): إن بناء و صياغة نموذج الحل السليم و الدقيق يتم من خلال عملية الخطأ و الصواب، حيث كثيراً ما تحتاج مثل هذه العملية إلى مساعدة المختصين في التنقيب عن البيانات بهدف اختبار و فحص مختلف البدائل للحصول على أفضل نموذج لحل المشكلة قيد الدراسة.

5- التقييم و تحليل نتائج النموذج (Evaluation and Interpretation): حالما يتم صياغة النموذج و التحقق من ثباته و صدقه، تجري مباشرة عملية التحقق من ثبات حزمة البيانات التي يتم تغذيتها بواسطة النموذج. وبما أن نتائج هذه البيانات معروفة، لذا فإن النتائج المتوقعة تقارن مع النتائج الفعلية في ثبات حزمة البيانات قيد التشغيل. وتؤدي هذه المقارنة أو المفاضلة إلى التحقق من دقة النموذج.

6- نشر و توزيع النموذج (Model Deployment) : حيث تشتمل هذه الخطوة على نشر و توزيع النموذج داخل المنظمة لمساعدة عملية صنع القرار. وأن النموذج الصالح يجب أيضاً أن يحقق الرضا لدى المستخدمين طالما أن اختيار النموذج لا بد أن يتم من خلال الدراسة الاسترشادية أو نموذج مصغر من الدراسة الشاملة.

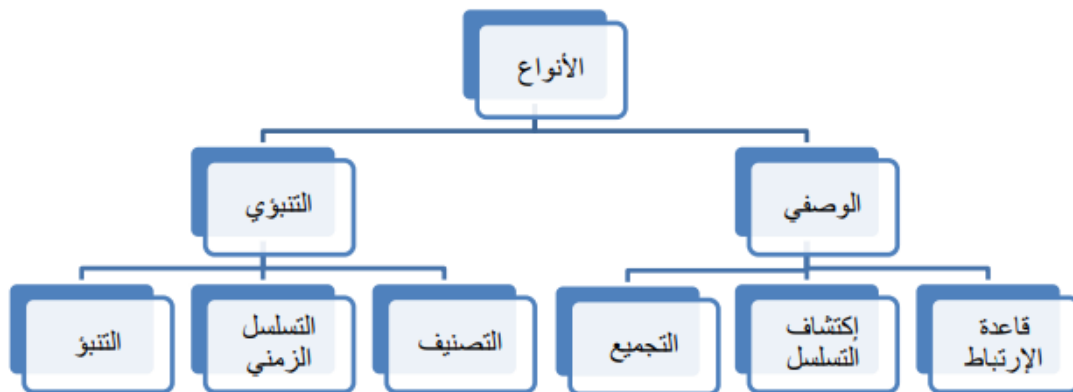


شكل (1) مراحل اكتشاف المعرفة⁵⁰

4-أنواع التنقيب في البيانات:

هناك نوعان أساسيا للتنقيب في البيانات هما(السيد،2015)⁵¹:

- 1- **التنقيب التنبؤي:** التنقيب التنبؤي ينتج عنه نموذج عن النظام الذي تصفه البيانات المستخدمة في التنقيب، ومن مهام التنقيب التنبؤي يقوم باستخدام بعض المتغيرات ليقوم بالتنبؤ ببعض القيم الغير معروفة أو المستقبلية لمتغيرات أخرى.
- 2- **التنقيب الوصفي:** ينتج عنه معلومات جديدة بناء على المعلومات الموجودة داخل البيانات المستخدمة في عملية التنقيب. من مهام التنقيب الوصفي إيجاد أنماط يمكن تفسيرها وفهمها بواسطة الانسان وتقوم هذه الأنماط بوصف البيانات.



شكل (2) يوضح أنواع تنقيب البيانات⁵²

G. Dong and J. Li. Efficient mining of emerging patterns: Discovering trends and differences. In Proc. 1999 Int. ⁵⁰ Conf. Knowledge Discovery and Data Mining (KDD'99), pp. 43–52, San Diego, CA, Aug. 1999.

⁵¹ سيد، أحمد فايز أحمد . دراسة بعنوان ادوات التنقيب عن البيانات مفتوحة المصدر(دراسة تحليلية تقييمية). (2015).

⁵² Fayyad, U. Piatetsky-Shapiro, G. & Smyth, P. (1996), —From data mining to knowledge Discovery in database||, American Association for Artificial Intelligence Press, Cambridge

5-أدوات التنقيب في البيانات:

في ظل كلاً من التنقيب التنبؤي، والتنقيب الوصفي، تستخدم عملية تنقيب البيانات أدوات عديدة تتمكن من خلالها من اكتشاف الاتجاهات والأنماط الخفية في حجم كبير من البيانات، والتي يمكن للمراجع الاعتماد عليها سواء مجتمعة أو بشكل منفرد في تطوير عملية المراجعة للوصول إلى الكفاءة والفاعلية المطلوبة وتشمل هذه الأدوات ما يلي(الدوري وآخرون، 2007)⁵³:

أولاً: أدوات التنقيب التنبؤي:

1- الانحدار (Regression): ويمكن الاعتماد عليه في تحليل 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. ويستخدم لتعيين عنصر البيانات إلى متغير التنبؤ بقيمة حقيقية. حيث تؤكد الدراسات أن الانحدار يعتبر أداة هامة يمكن أن تساعد المراجع في كشف الاحتيال وتحديد الأخطاء.

2- التصنيف (Classification): ومن خلاله يتم تحليل مجموعة من البيانات لتكوين مجموعة من

القواعد المتجمعة التي يمكن أن تستخدم لتصنيف بيانات المستقبل، أي إيجاد المعلومات التي تتعلق بالخصائص المشتركة. ويمكن الاعتماد على التصنيف في تحقيق عدة مهام مثل دراسة اتجاهات الربح المستقبلي، كما يمكن فحص عمليات الائتمان التي قامت بها المنشأة لاستنتاج العمليات المزورة أو الوهمية.

3- التنبؤ (Prediction): يُعد التنبؤ من الأدوات التي تجذب الانتباه لأنها تتمكن من إعطاء مغزى

للتوقع الناجح في سياق العمل لذا فإنه يمكن النظر إلى العديد من تطبيقات التنقيب كأنها تنبؤ بحالة بيانات مستقبلية معتمدة على بيانات سابقة وحالية.

ويشبه التنبؤ التصنيف أو التقدير، ما عدا أن البيانات تصنف على أساس التنبؤ بسلوكها المستقبلي أو تقدير قيمتها المستقبلية. حيث أن المتغير التابع المتنبئ به هو متغير كمي. ومن خلال التنبؤ يمكن دراسة سلوك وموثوقية العملاء، مما يساعد في وضع نسب الديون المشكوك فيها بطريقة أقرب ما تكون للواقع وبالتالي يصبح رقم مخصص الديون المشكوك فيها مناسب وبما ينعكس على تحديد قيمة الربح في قائمة الدخل وأيضاً إظهار المركز المالي السليم للشركة، وهو ما يؤدي إلى تحقيق الأهداف المرجوة من عملية المراجعة.

⁵³ د. زكريا مطلق الدوري، د. داليا عبد الحسين احمد (٢٠٠٧)، دور تنقيب البيانات في زيادة أداء المنظمة - دراسة تحليلية في المصرف الصناعي، مجلة العلوم الاقتصادية والإدارية، جامعة بغداد، العدد (٤٨)، المجلد (١٣)، ص ٤٠ - ٦٢

ثانياً: أدوات التنقيب الوصفي:

1- **الاقتران (Association):** وهو يتضمن علاقات اقتران ثابتة بين مجموعة من الأشياء في قاعدة البيانات. أي الاقتران بين حدوث حدث ما وحدث حدث آخر. فمثلا إعلان إفلاس أحد العملاء لابد أن يؤدي إلى زيادة رقم الديون المعدومة ومن ثم يساعد الاقتران المراجع في إظهار مثل هذه العلاقات.

2- **التجميع (Clustering):** وهو عملية تجميع السجلات المتشابهة في مجموعات، ويتم ذلك بهدف الاستكشاف عالي المستوى لما يجري داخل قاعدة البيانات .

فهو طريقة لتجزئة البيانات بشكل عنقودي ويتم تجميعها على شكل فئات عنقودية مختلفة لدراستها بطريقة يمكن أن تساهم في تفسيرها واستنباط ما فيها من معرفة.

3- **التلخيص (Summarization):** ينطوي التلخيص على إيجاد وصفا لمجموعة فرعية من البيانات المدمجة. وغالباً ما تطبق تقنيات التلخيص لتحليل البيانات التفاعلية الاستكشافية وتوليد تقرير آلي.

وتعتمد الأدوات السابقة على عدة أساليب فرعية يساعد كل منها في تحقيق هدفاً معيناً، ومن أهمها ما يلي:

أ- **خوارزمية الجار الأقرب (Nearest Neighbor):** وتعتبر من تقنيات التنقيب في البيانات التي تهدف للتنبؤ عن طريق مقارنة السجلات الشبيهة بالسجل المراد التنبؤ له وتقدير القيمة المجهولة لهذا السجل بناء على معلومات لتلك السجلات (الطيار وآخرون، 2005)⁵⁴. ويمكن استخدام الجار الأقرب في الكشف عن محاولات الغش والتدليس في العمليات المالية من خلال استخدام المعلومات الخاصة بالمحاولات أو الأحداث السابقة والتي تمت في المنشأة بهدف الغش أو التدليس لتحديد أو اكتشاف المحاولات الحالية عن طريق تحديد عدد من السجلات التجريبية ثم استخدامها بهدف التنبؤ بالقيمة المطلوبة. إن الفكرة الأساسية لخوارزمية الجار الأقرب تكمن في تصنيف الحالات غير المرئية أو (غير المصنفة) إلى الحالات الأقرب لها ضمن حجم معين. وهو ما يساعد في إيجاد حلول لمشكلات جديدة من خلال ملاحظة مشكلات سابقة تم حلها. ويطبق المراجع الجار الأقرب بدرجة ثقة معينة، ويعبر عن ذلك بأن يقول إنه على ثقة ٧٠ % مثلا من قيمة معينة قام باستكشافها، ويتم تحديد درجة الثقة بناء على ما يلي (الطائي و آخرون، 2011):

- المسافة بين السجل المستكشف وأقرب جار.

- مدى تجانس مجموعة الجوار وما إذا كانت تؤدي لنفس القيمة المستكشفة.

⁵⁴ طيار، أحسن. شلابي، عمار (٢٠٠٥)، التنقيب في البيانات واتخاذ القرارات، مجلة كلية العلوم الاقتصادية وعلوم التسيير، جامعة ٢٠ أوت ١٩٥٥ سكيكدة، الجزائر، العدد الرابع، ص ٨٤ - ٩٤.

ب- **شجرة القرارات (Decision Trees):** وهي هيكل على شكل شجرة الذي يمثل مجموعة من القرارات. هذه القرارات تولد قواعد لتصنيف مجموعة البيانات، فشجرة القرار هي أسلوب استكشافي يظهر على شكل شجرة، وبشكل دقيق يمثل كل فرع من فروعها سؤالاً تصنيفياً وتمثل أوراقها أجزاءً من قاعدة البيانات تنتمي للتصنيفات التي تم بناؤها (عبد الله، 2004) 55.

والقاعدة الأساسية في بناء شجرة القرارات هي إيجاد أفضل سؤال عند كل فرع من فروع الشجرة بحيث يقسم هذا السؤال البيانات إلى قسمين، القسم الأول منها ينطبق عليهم السؤال والقسم الثاني لا ينطبق. وهكذا يتم من خلال سلسلة من الأسئلة بناء شجرة القرار بفروعها المتسلسلة (بابكر، 2007) 56.

وبالرغم من أن شجرة القرارات تستخدم في الاستكشاف وتحضير البيانات للعمليات الإحصائية إلا أنها أيضاً تستخدم وبشكل أكثر للتنبؤ. ومن المهم عند بناء خوارزمية شجرة القرار أن يؤخذ بعين الاعتبار أن تكون قابلة للتطبيق بقدر الإمكان وبشكل مثالي على كل البيانات المتوفرة (Barry de Ville, 2006) 57.

وتعتبر شجرة القرارات من الأدوات التي يمكن أن يعتمد عليها في حل المشكلات، خاصة في حالة أن يمر حل المشكلة بعدة مراحل، كما أن شجرة القرارات تساعد على استخدام الاحتمالات المشتركة واللاحقة للتوصل إلى أفضل حل للمشكلة.

ج- **الشبكات العصبية (Networks Neural):** تعد الشبكات العصبية الاصطناعية أهم أنواع الذكاء الاصطناعي، وتتمحور فكرتها حول محاكاة قدرة العقل البشري على التعرف على الأنماط وتمييز الأشياء، باستخدام الحاسب من خلال إتباع عملية التعلم الذاتية التي تحدث في العقل، والتي يتم فيها الاستفادة من الخبرات السابقة في سبيل الوصول إلى أفضل نتائج في المستقبل.

ويعمل هذا الأسلوب في ظل ظروف مناسبة للتنبؤ بظروف الحال المستقبلية وبدون سابق افتراض أو استخدام أية نماذج افتراضية لتمثيل البيانات، الأمر الذي لا يتطلب ضرورة التوصيف المسبق للفروض بين العلاقات المختلفة، بل أنه سوف يحدد هذه العلاقات من خلال عملية تعلم خاصة به، تمر في عدة مراحل من التنقيح والترشيح الذي يتناسب في النهاية مع هذه البيانات، وعلاقة المتغيرات فيما بينها، يضاف إلى ما تقدم، فإن درجة الثقة التي يحتاجها النموذج لاختبار مصداقية النتائج تعتبر أصغر

55 عبد الله، خالد أحمد. (٢٠٠٤)، "قواعد البيانات المتقدمة - مستودعات البيانات"، كلية علوم الحاسوب وتقانة المعلومات، جامعة العلوم والتقانة، السودان، يناير، ص ٢٠١.

56 بابكر، مصطفى. (٢٠٠٧)، التحليل باستخدام شجرة القرار، المعهد العربي للتخطيط، الكويت، ص ١ - ١٨.

57 Barry de Ville (2006), Decision Trees for Business Intelligence and Data Mining: Using SAS Enterprise Miner, SAS Institute Inc., Cary, NC, USA, 2006, p1.

نسبياً من تلك التي تحتاج إليها النماذج الإحصائية التقليدية، الأمر الذي يبرز أهمية ذلك الأسلوب في التنبؤ، وبما يمكن لاحقاً من تنقيح التوقعات والتنبؤات (Kotsiantis et al,2006,p672-681)⁵⁸. ويمكن الاعتماد على الشبكات العصبية في تحقيق عدة مهام منها (جمعة،2012،ص 479-505)⁵⁹:

- التنبؤ بالعوائد المتوقعة.
- تحليل فرص استمرارية المنشأة.
- اكتشاف الخطأ، والتنبؤ بحالات الغش.

د- نظرية بيز المبسط (Naïve Bayes): يتم استخدامها كمصنف احتمالي بسيط يعتمد على قاعدة بيز للاحتمال الشرطي. تعتمد نظرية بيز المبسط على افتراضات إحصائية قوية لمتغيرات التنبؤ المستقلة. وأحد استخداماتها هو "فلتر للبريد المزعج" Email Spam والتي تستخدمه شركات مثل Hotmail و Yahoo كما تستخدم هذه الطريقة بشكل كبير في البنوك وفي اكتشاف الغش المالي (عبيدالله،2016،ص6)⁶⁰.

هـ- الخوارزميات الجينية (Algorithm Genetic) و منطق الغموض (Logic Fuzzy): تستخدم الخوارزميات الجينية لأغراض التصنيف لتمثل سلوك متخذ القرار في حالات مثل الغش، كما تستخدم في حالات حل مشاكل الخروج عن المثالية مثل مشاكل تزوير بطاقات الائتمان غير المحددة بشكل جيد، كما يمكن استخدام الخوارزميات الجينية في تقييم مدى ملائمة الأساليب الأخرى للتنقيب عن البيانات. أما منطق الغموض فهو أسلوب إحصائي يصنف الأسباب الحكيمة ويخصص البيانات لمجموعة أو فئة معينة استناداً إلى درجة إمكانية وجود هذه البيانات في هذه المجموعة (عبيدالله،2016،ص6)⁶¹.

وتجدر الإشارة إلى أن هذه الأساليب لا تُستخدم بشكل فردي وإنما مع كل استراتيجيات للتنقيب في البيانات (مثل التصنيف والتنبؤ وتحليل التجميع واكتشاف العلاقات) يمكن استخدام عدة أساليب علمية للتنقيب في البيانات والكشف عن أنماط وعلاقات جديدة، فعلى سبيل المثال من أشهر الأساليب المستخدمة مع

⁵⁸ S. Kotsiantis , E. Koumanakos , D. Tzelepis , V. Tampakas (2006), Financial Application of Neural Networks: Two Case Studies in Greece, Artificial Neural Networks - ICANN, 16th International Conference, Athens, Greece, September 10-14, PP. 672 – 681.

⁵⁹ جمعة، أحمد حلمي (٢٠١٢)، استخدام الشبكات الاصطناعية في اكتشاف الأخطاء الجوهرية في البيانات المالية: دراسة تطبيقية، بحث مقدم إلى المؤتمر العلمي السنوي الحادي عشر بعنوان "ذكاء الأعمال واقتصاد المعرفة"، كلية الاقتصاد والعلوم الإدارية، جامعة الزيتونة الأردنية ٢٣ – ٢٦ ابريل، ص ص. ٤٧٩ – ٥٠٥

⁶⁰ عبيد الله،فايزة.(2016)، دعم أساليب القياس واتخاذ القرار باستخدام أساليب التنقيب عن البيانات كمدخل لإدارة التكاليف الاستراتيجية و دعم القدرة التنافسية للشركات الصناعية، كلية التجارة، جامعة دمنهور، دورية، ص 9.

⁶¹ عبيد الله،فايزة.(2016)، دعم أساليب القياس واتخاذ القرار باستخدام أساليب التنقيب عن البيانات كمدخل لإدارة التكاليف الاستراتيجية و دعم القدرة التنافسية للشركات الصناعية، كلية التجارة، جامعة دمنهور، دورية، ص 9.

استراتيجية التصنيف هي الشبكات العصبية ونظرية ببيز المبسط وشجرة القرارات والخوارزميات الجينية، حيث تستطيع هذه الأساليب تكوين نموذج من بيانات معروف تصنيفها (Koksal et al,2011)⁶². ومن أشهر الأساليب استخداماً مع استراتيجية التنبؤ هي الشبكات العصبية ونموذج التنبؤ اللوجستي وشجرة القرارات، حيث تساعد هذه الأساليب في استخدام النماذج الناتجة عن استراتيجية التصنيف للتنبؤ بالسلوك المستقبلي في مجتمع تصنيفه غير معروف (Sharma et al,2012)⁶³. وكذلك من أشهر الأساليب المستخدمة مع استراتيجية تحليل التجميع هي تحليل الجار الأقرب ونظرية ببيز المبسط، حيث تُستخدم هذه الأساليب مع استراتيجية تحليل التجميع لاستنتاج عناوين التصنيف الناتجة من استراتيجية التصنيف. وذلك بهدف تكوين مجموعات متشابهة وذات معنى ولم تكن معروفة سابقاً في شكل نماذج (Koksal et al,2011)⁶⁴.

6-تطبيقات التنقيب في البيانات:

يلعب التنقيب عن البيانات دوراً أساسياً في البنوك، والتأمين والنقل والتجارة، ويمكن للتنقيب عن البيانات حل كثير من المشاكل الحسابية، وزيادة الأرباح وصنع قرارات حكيمة. ولم تكن العمليات التجارية التطبيق الأولى لتقنيات التنقيب عن البيانات إنما كانت من المجالات المهمة، لأن العمليات التجارية بها الكثير من بيانات المبيعات، مثل تسجيلات الشراء الخاصة بالمستهلكين، ومعلومات المستهلكين، ومعلومات الخدمة وغيرها. يمكن للشركات استخدام البيانات لتصنيف المستهلكين من مجموعات المستهلكين الأساسية وإيجاد الخصائص المشتركة للمستهلكين و رغباتهم المستقبلية، وتقديم منتجات كافية وخدمات تلبي رغباتهم. وعند استخدام تطبيقات برامج التنقيب عن البيانات، يجب اختيار الخوارزمية المناسبة، والمعرفة فيما وراء البيانات التي يمكن إيجادها. وبما أن التنقيب عن البيانات يقوم بعدة مهام مثل: تجميع البيانات وتخزينها وتنظيمها لذا فهو يستخدم في عدة مجالات مثل الطب والمالية والذكاء الصناعي والقانون والدفاع والتعليم وعمليات التحكم وغيرها، وتستخدم معظم التطبيقات التنقيب عن البيانات للدعاية والتسويق والمبيعات، كما يمكن استخدامه في التشخيص (السيد، 2015)⁶⁵.

Koksal, G., Batmaz, I., & Testik, M. C. (2011). A Review of Data Mining Applications for Quality Improvement in Manufacturing Industry. Expert Systems with Applications 38, pp. 13448-13467.

Sharma, A. & Panigrahi, P. K. (February 2012). A Review of Financial Accounting Fraud Detection Based on Data Mining Techniques. International Journal of Computer Applications 39 (1), pp. 37-47.

Koksal, G., Batmaz, I., & Testik, M. C. (2011). A Review of Data Mining Applications for Quality Improvement in Manufacturing Industry. Expert Systems with Applications 38, pp. 13448-13467.

65 سيد، أحمد فايز أحمد . دراسة بعنوان ادوات التنقيب عن البيانات مفتوحة المصدر (دراسة تحليلية تقييمية). (2015).

1.6. تطبيقات التنقيب في البيانات في القطاع المصرفي:

شهدت الصناعات المصرفية في جميع أنحاء العالم تغيرات هائلة في طريقة إدارة الأعمال. ومع ازدياد استخدام الخدمات المصرفية الالكترونية، أصبح الحصول على بيانات المعاملات أسهل، ونتيجة لازدياد حجم البيانات بشكل كبير أصبح من المهم تحليل هذه البيانات وتحويلها الى معرفة مفيدة للمؤسسة.

وباستخدام تقنيات التنقيب في البيانات لتحليل الاتجاهات والأنماط يمكن لمتخذي القرار في البنوك التنبؤ بتفاعل العملاء مع تعديلات أسعار الفائدة، وتقبل العملاء لعروض المنتجات الجديدة، والتنبؤ بالعملاء المعرضين لمخاطر التخلف عن السداد، وكيفية جعل العلاقات مع العملاء أكثر ربحية. ونذكر بعض تطبيقات التنقيب في البيانات في القطاع المصرفية (madanLalBhasin,2006)⁶⁶:

1- التسويق: يعتبر التسويق هو أحد أكثر مجالات التنقيب في البيانات استخداماً في القطاع المصرفي، يمكن لقسم التسويق في البنك استخدام التنقيب في البيانات لتحليل قواعد بيانات العملاء لتمييز بين العملاء المربحين وغير المربحين، حيث يقوم التنقيب في البيانات بتحليلات مختلفة من البيانات المخزنة لتحديد سلوك المستهلك من خلال الرجوع الى المنتج والسعر والتوزيع، ويمكن معرفة رد فعل العملاء على المنتجات الحالية والجديدة، والتي تساعد المصارف على الترويج للمنتج وتحسين جودة المنتجات والخدمات واكتساب ميزة تنافسية، ويمكن لمحللي البيانات تحليل الاتجاهات السابقة وتحديد الطلب الحالي والتنبؤ بسلوك العملاء في مختلف المنتجات والخدمات (Desai et al,2004)⁶⁷.

2- إدارة المخاطر: يستخدم التنقيب في البيانات على نطاق واسع لإدارة المخاطر في القطاعات المصرفية. ويحتاج المسؤولون التنفيذيون الى معرفة فيما إذا كان العملاء الذين يتعاملون معهم جديرين بالثقة أم لا، وإن قرار توسيع حدود الائتمان للعملاء الحاليين، ومنح ائتمان لعملاء جدد، والموافقة على القروض جميعها قرارات محاطة بالمخاطر بالنسبة للبنوك إذا لم يكن لديها معلومات مسبقة عن عملائها.

وعند قيام البنك بعملية المنح الائتماني يقوم أولاً بالتحقق من التفاصيل المختلفة المتعلقة بالقرض، ك مبلغ القرض، معدل الفائدة، فترة السداد، نوع الضمانة البيانات الديموغرافية المتعلقة بطالب القرض، والتاريخ الائتماني للمقترض. والعملاء الذي لديهم تاريخ طويل في التعاملات الائتمانية مع المصارف ومن ذوي

⁶⁶ MadanLalBhasin, —Data Mining: A Competitive Tool in the Banking and Retail Industries||, The Chartered Accountant October 2006

⁶⁷ B. Desai and Anita Desai, "The Role of Data mining in Banking Sector", IBA Bulletin, 2004.

الدخل المرتفع لديهم فرص أكبر في الحصول على قروض بسهولة دون غيرهم. وعلى الرغم من أن البنوك تكون حذرة بعملية المنح الائتماني وأثناء تقديم القروض ولكن هناك دائماً فرصاً للتخلف عن السداد. وهنا يكون دور التنقيب في البيانات في مساعدة متخذي القرار بتحليل سلوك وموثوقية عملاء الائتمان والتميز بين المقترضين الجيدين والملتزمين بالسداد عن غيرهم، وبهذا يستطيع البنك من تقليل من المخاطر المحتمل حدوثها (madanLalBhasin,2006)⁶⁸.

3- الكشف عن الاحتيال: تعد القدرة على اكتشاف الإجراءات الاحتيالية مصدر قلق متزايد للعديد من الشركات والمؤسسات المالية، وبمساعدة التنقيب في البيانات يتم الكشف عن المزيد من الإجراءات الاحتيالية والإبلاغ عنها. حيث تم تطوير نهجين مختلفين من قبل المؤسسات المالية لاكتشاف أنماط الاحتيال، في النهج الأول يبحث البنك في مستودع البيانات الخاص بطرف ثالث ويستخدم برامج التنقيب في البيانات لتحديد أنماط الاحتيال، ويمكن بعد ذلك إحالة هذه الأنماط الى قاعدة البيانات الخاصة به بحثاً عن إشارات تدل على وجود مشكلة داخلية. وفي النهج الثاني يعتمد تحديد نمط الاحتيال على المعلومات الداخلية الخاصة بالبنك، ومعظم البنوك تستخدم نهجاً هجيناً لتحليل البيانات والمعلومات الداخلية (Kaptan et al,2002)⁶⁹.

4- إدارة العلاقات مع العملاء: يعد اكتساب العملاء والاحتفاظ بهم من الاهتمامات المهمة جداً لدى أي قطاع وبشكل خاص القطاع المصرفي (madanLalBhasin,2006)⁷⁰.

يمتلك العميل اليوم خيارات واسعة جداً من الخدمات والمنتجات التي تقدمها البنوك المختلفة، وهنا يتعين على البنك تلبية احتياجات العملاء من خلال توفير هذه المنتجات والخدمات التي يفضلونها والتي بدوره يؤدي الى ولاء العملاء والاحتفاظ بهم، وهنا يكون دور التنقيب في البيانات في تحليل العملاء المخلصين ممن ينتقلون الى بنوك أخرى للحصول على خدمات أفضل، وبالتالي معرفة أسباب هذا التحويل ومعالجتها مما يساعد البنوك على الأداء بشكل أفضل واكتساب ولاء عملائها (Desai et al,2004)⁷¹.

⁶⁸ MadanLalBhasin, —Data Mining: A Competitive Tool in the Banking and Retail Industries||, The Chartered Accountant October 2006

⁶⁹ S. S. Kaptan, N S Chobey, "Indian Banking in Electronic Era", Sarup and Sons, Edition 2002.

⁷⁰ MadanLalBhasin, —Data Mining: A Competitive Tool in the Banking and Retail Industries||, The Chartered Accountant October 2006

⁷¹ B. Desai and Anita Desai, "The Role of Data mining in Banking Sector", IBA Bulletin, 2004.

7- استراتيجيات نجاح التنقيب في البيانات:

لكي يتم الحصول على تنقيب ناجح، لابد من اتباع عدد من الاستراتيجيات التي تتضمن ذلك ومنها (الدوري وآخرون، 2007، ص40-62)⁷²:

- 1- توجيه العمل نحو هدف محدد ويوضح المشاكل المراد حلها.
- 2- توفير بيانات ذات كمية ومحتوى عال يزيد من قوة التنبؤ، فلما كانت البيانات أفضل، كانت النتائج أفضل أي ذات تحليل شامل وأكثر دقة.
- 3- بناء عدة نماذج باستخدام عينات من البيانات بدلاً من بناء نموذج واحد باستخدام كل البيانات.
- 4- التهيئة الصحيحة للبيانات والإسراع في استخدام تنقيب البيانات تساهم بشكل كبير وانجاح تنقيب البيانات.
- 5- التواصل المستمر بين مدراء العمل والخبراء مع إمكانية احتواء متخذ القرار من أجل إعادة تعريف أو تصميم أهداف العمل وإيجاد النتائج الأولية وتكرار ذلك مع ازدياد إدراك البيانات.
- 6- التخطيط للتعلم من عملية تنقيب البيانات وتعليمها الى من له صلة مباشرة باستخدامها في اتخاذ القرارات، وعدم اهمال أي نتائج تحت ضغط الوقت والكلفة.
- 7- التكامل بين عدد من الأدوات بضمنها القديمة والجديدة.

8- صعوبات التنقيب في البيانات:

اكتشاف المعرفة والتنقيب في البيانات محورين جديدين من محاور المعلوماتية التي تتعامل مع معطيات متنامية بوتيرة عالية، وما زال هذا العلم يواجه الكثير من الصعوبات منها(بسيوني، 2017)⁷³:

- 1- عملية اكتشاف المعرفة: هناك حاجة الى فهم أفضل لعملية اكتشاف المعرفة، إذ يجب تحديد المراحل المختلفة في هذه العملية، والعلاقات بين مختلف الخطوات، كذلك ينبغي وضع دليل يساعد محلي الأعمال ومهندسي التنقيب في انجاز مشاريع اكتشاف المعرفة.
- 2- التغيرات والتبدلات في البيانات: إن تبدل البيانات باستمرار، وحفظها في قواعد البيانات يجعل من عملية اكتشاف المعرفة محض الإثارة لأولئك الذين يسعون الى إيجاد تقانات جديدة للتعامل مع تلك الأشكال المتغيرة، مما يضطرهم الى إيجاد وسائل إضافية للتأكد من أن عملية اكتشاف المعرفة

⁷² د. زكريا مطلق الدوري، د. داليا عبد الحسين احمد (٢٠٠٧)، دور تنقيب البيانات في زيادة أداء المنظمة - دراسة تحليلية في المصرف الصناعي، مجلة العلوم الاقتصادية والإدارية، جامعة بغداد، العدد (٤٨)، المجلد (١٣)، ص٤٠ - ٦٢

⁷³ بسيوني، عبد الحميد. (2017)، الحاسب واكتشاف المعرفة تنقيب البيانات، دار الكتب العلمية للنشر والتوزيع، القاهرة.

المستخدمة مازالت فعّالة لدى الانتقال الى صيغ جديدة للمعطيات، وفي بعض الأحيان تعديل النموذج المستخدم ليتناسب معها.

3- المحاور المتعددة الاستراتيجية: أحياناً يكون استخدام نوع واحد من خوارزميات التنقيب في البيانات غير كافٍ للإجابة عن سؤال معقد نوعاً ما ولذلك تصبح الحاجة الى توظيف عدة استراتيجيات قائمة على استخدام مجموعة من التقانات للإجابة عن سؤال واحد.

4- تدرج عملية اكتشاف المعرفة والتنقيب في البيانات: يزداد كمّ البيانات بسرعة، وتتطلب قواعد البيانات اليوم استخدام طرق فعّالة للوصول الى المعطيات. وهي أيضاً مجبرة على استخدام خوارزميات التنقيب في البيانات، التي يجب تعديلها بما يتناسب مع هذا الكمّ المتزايد من البيانات. كذلك ينبغي توظيف تقانات لتلخيص المعطيات قبل القيام بعملية التنقيب فيها ضمن عملية اكتشاف المعرفة.

أما تحديات التنقيب في البيانات تتمثل فيما يلي (بسيوني، 2017)⁷⁴:

- 1- القياسية Scalability.
- 2- مستويات الأبعاد Dimensionality.
- 3- البيانات المعقدة وغير المنتظمة Complex and Heterogeneous Data.
- 4- جودة البيانات Data Quality.
- 5- ملكية البيانات والتوزيع Data Ownership and Distribution.
- 6- الحفاظ على الخصوصية Privacy Preservation.
- 7- تدفق البيانات Streaming Data.

ثالثاً: دور التنقيب في البيانات في اتخاذ قرارات منح الائتمان:

تكتسب القرارات الائتمانية أهمية كبيرة لعلاقة التشابك والتداخل بين هذه القرارات ونتائجها وبين الاستراتيجية العامة للبنك واستراتيجيته التسويقية على وجه الخصوص.

واتخاذ قرارات الائتمان ينبغي أن تتوافر لديها مجموعة من الخصائص والمقومات الهامة التي تمكنها من القيام بعملها بكفاءة وفعالية وأن تتوافر لديها خبرات متنوعة ومعارف متعددة وصفات خاصة تؤهلها للقيام بهذه المهمة الصعبة.

⁷⁴ بسيوني، عبد الحميد. (2017)، الحاسب واكتشاف المعرفة وتنقيب البيانات، دار الكتب العلمية للنشر والتوزيع، القاهرة.

وكما يعد اتخاذ القرار الائتماني الخاص بمنح التسهيلات الائتمانية أو رفض منحها من أخطر القرارات التي تتخذ في البنوك، فاتخاذ قراراً خاطئاً في هذا المجال له آثار سلبية وخيمة على أوضاع البنك، بل ربما على مستقبله، ولذلك يجب محاولة اتخاذ قرارات مثلى للمحافظة على المركز المالي للبنك (لعروسي القرين، 2017)⁷⁵.

1- خصائص القرار الائتماني:

تتمحور خصائص القرار الائتماني والتي وإن توفرت يصعب حدوث التعثر بخلاف ما يرجع الى أسباب يصعب السيطرة عليها وتقع خارج دائرة مسؤولية العميل والبنك (الشواربي، 2002)⁷⁶:

- 1- **خاصية السيولة في الائتمان الممنوح:** وهي تعني إمكانية التصفية الذاتية للائتمان وسداده كاملاً في التاريخ المتفق عليه وبالشروط المتفق عليها، لذا يتعين على الفاحص الائتماني التحقق من ذلك بالإضافة الى سلامة استخدام التسهيل في الغرض الممنوح من أجله والتحقق من مصدر واضح للسداد.
- 2- **الربحية المتوقعة من الائتمان الممنوح:** إذ يجب الموازنة بين المخاطر المدروسة المصاحبة للقرار الائتماني وبين الربحية المتوقعة منه في إطار السياسة الائتمانية للبنك فيما يتعلق بنوع التسهيلات الممكن تقديمها وأسعار العائد واجبة التطبيق.
- 3- **خاصية الأمان في الأموال المقرضة:** وتعني الثقة في أن الأموال المقرضة سوف يتم سدادها في تاريخ الاستحقاق ويأتي ذلك من الحرص الشديد على توافر عامل السيولة بالإضافة الى عدم السماح بمنح ائتمان لأغراض المضاربة أو لأغراض مظهرية.
- 4- **خاصية تنوع محفظة القروض والتسهيلات:** والهدف من تنوع محفظة القروض والتسهيلات بالبنك هو توزيع المخاطر بتجنب مخاطر التركيز على منطقة جغرافية معينة أو نشاط أو قطاع اقتصادي معين أو عميل بعينه أو التركيز على نوع وحيد من الضمانات لأن أية تقلبات غير متوقعة سوف تؤثر سلباً على المركز المالي للبنك.
- 5- **خاصية الضمان:** والذي يعتبر خط دفاع أخير لحالات الطوارئ غير المنظورة أو لمجابهة الحالات التي تحيط بها مخاطر ودرجة عالية من التأكد، أخذاً في الاعتبار أن خاصية الضمان تأتي في المرتبة الأخيرة في منظومة القرار الائتماني الأمثل.

⁷⁵ زهرة، لعروسي قرين. (2017): "دور إدارة مخاطر الائتمان المصرفي في اتخاذ القرارات الائتمانية لدى البنوك التجارية-دراسة مجموعة من البنوك التجارية الجزائرية". رسالة دكتوراه منشورة، كلية العلوم الاقتصادية والتجارية علوم التسيير، جامعة محمد بوضياف بالمسيلة.
⁷⁶ الشواربي، عبد الحميد محمد. إدارة المخاطر الائتمانية من جهتي النظر المصرفية والقانونية، منشأة المعارف، الإسكندرية، مصر، 2002، ص552-553.

2-مراحل صناعة القرار الائتماني:

إن صناعة قرار الائتمان أي اتخاذ القرار بمنح الائتمان يمر بثلاث مراحل هي كالتالي (غنيم، 2002):⁷⁷

1- مرحلة ما قبل اتخاذ القرار الائتماني: وتعرف بالمرحلة التمهيديّة لصناعة القرار الائتماني وفيها يتم:

- النشاط الترويجي للخدمات التي يمكن للبنك تقديمها.
- الاستعلامات.
- التناول المبدئي لطلب الاقتراض.
- التفاوض مع طالب القرض على الشروط المقترحة لمنحه الائتمان.

2- مرحلة اتخاذ القرار الائتماني: وتعرف بمرحلة تخليق القرار الائتماني وفيها يتم:

- تحديد المخاطر الائتمانية لطلب القرض.
- ضوابط مواجهة هذه المخاطر.
- قياس المخاطر الائتمانية.
- تقييم الجدارة الائتمانية لطالب القرض.
- صياغة القرار الائتماني.
- اعتماد السلطة المختصة للقرار الائتماني.

3- مرحلة ما بعد اتخاذ القرار الائتماني: وهي مرحلة ترجمة القرار (اختباره) عملياً وفيها يتم:

- السيطرة على الضمانات.
- استخدام التسهيلات الائتمانية.
- المتابعة المكتبية والميدانية لاستخدام التسهيلات في الغرض الممنوحة له.
- سداد التسهيلات التي تم منحها.

3-مقومات القرار الائتماني والعوامل المؤثرة فيه:

يتأثر القرار الائتماني بعوامل متعددة منها ما يتعلق بالمصرف ومنها ما يتعلق بالعميل كما أنه هناك مقومات

لقرار ائتماني أمثل (الشواربي، 2002):⁷⁸

77 غنيم، أحمد. صناعة قرارات الائتمان والتمويل في إطار الاستراتيجية الشاملة للبنك، بدون ذكر دار النشر، بدون ذكر البلد، 2002، ص 67.
78 لشواربي، عبد الحميد محمد. إدارة المخاطر الائتمانية من وجهتي النظر المصرفية والقانونية، منشأة المعارف، الإسكندرية، مصر، 2002، ص 552-553.

1- مقومات القرار الائتماني: يتطلب القرار الائتماني الأمثل عدداً من المقومات يمكن إجمالها في النقاط التالية:

- أن يمنح الائتمان لصالح عميل تتوافر له الجدارة الائتمانية وفق معايير هذه الجدارة المتعارف عليها.
- أن يكون الائتمان موجهاً لغرض واضح ومشروع يتفق مع طبيعة نشاط العميل.
- تناسب قيمة الائتمان الممنوح من الغرض الصادر من أجله والمركز المالي والهيكل التمويلي للعميل المقترض.
- التأكد من مصدر سداد واضح ومحدد ومرتبطة مباشرة بغرض استخدام الائتمان وتحيط به درجات عالية من الثقة والتأكد، ليس فقط أثناء فترة منح الائتمان ولكن من خلال سريان هذا الائتمان ولحين وقوع تاريخ السداد المتفق عليه.
- التأكد من أن الذي يتصدى لصناعة قرار منح الائتمان كوادر وسلطة ائتمانية تتوافر لهما المقومات الشخصية والموضوعية اللازمة.

2- العوامل المؤثرة على القرار الائتماني: توجد مجموعة من العوامل التي تؤثر على القرار الائتماني نذكر منها:

- التخطيط الاستراتيجي للبنك (الأهداف التي يسعى البنك لتحقيقها).
- الاستراتيجية المصرفية التي يطبقها البنك لتحقيق أهدافه العامة.
- الاستراتيجية والخطة التسويقية للبنك.
- ضوابط ومحددات السياسة الائتمانية للبنك.
- هيكل الموارد المالية للبنك (عناصره، تكلفته، درجة استقراره).
- هيكل التكاليف بالبنك.
- الكوادر البشرية المؤهلة للعمل داخل البنك.
- كيفية صناعة القرار الائتماني داخل البنك.
- الموقف التنافسي للبنك في السوق المصرفي.
- سياسة البنوك المنافسة وردود أفعالها داخل السوق.
- المناخ الاقتصادي والاجتماعي والسياسي العام للمجتمع.
- درجة النمو الاقتصادي للمجتمع.

ومنه نستنتج أن من أهم النشاطات التي يؤديها البنك هو عملية منح الائتمان، حيث تقدير الجدارة الائتمانية للعميل الأساس الذي تستند عليه الكثير من البنوك العالمية في عملية منح الائتمان المصرفي. فكلما كانت التقديرات دقيقة حقق المصرف هدفه الأساسي المتمثل باستمرار الأموال وبعكسه يتعرض المصرف الى

المخاطر الائتمانية فيما لو كانت التقديرات غير صحيحة وغير موضوعية، لذا يعد تقدير الجدارة الائتمانية للزبون المحصلة النهائية لعملية جمع وتحليل وتقويم المعلومات الائتمانية. إذ ينبغ إداء وتقويم الأركان الأساسية للجدارة الائتمانية، وبالتالي تكوين صورة متكاملة للزبون والتنبؤ بمستقبله العام من خلال عملية اسقاط لوضعه الائتمان في الماضي والحاضر.

وكما يعد تحليل المعلومات العامل المساعد للإدارة في اتخاذ القرارات الخاصة بمنح الائتمان، مما له من أهمية في التعرف على قابلية المقترض ورغبته وقدرته على التسديد وفق الشروط التي تم الاتفاق عليها مع المصرف وينبغي على المصرف تقدير حجم المخاطر التي يتحملها نتيجة تقديم القروض، ويعتمد ذلك على دقة تحليل المعلومات الائتمانية الذي يكون متفاوتاً من حيث الوقت والجهد بين القروض المختلفة ويعتمد ذلك على مقدار مبلغ القرض والغرض منه والضمانات المقدمة في ذلك.

ومن هنا يأتي دور التنقيب في البيانات في دعم واتخاذ القرارات الائتمانية، حيث أصبح موضوع التنقيب عن البيانات من المفاهيم الحديثة في نجاح نظم دعم القرار في منظمات الأعمال الحديثة (Linoff et al,2000)⁷⁹، ولغايات توفير معلومات تساعد على اتخاذ القرارات الجيدة (Dilnutt,2002)⁸⁰. وبرزت الحاجة الى مرحلة إجرائية لاستخلاص مواصفات وعلاقات من المعطيات، وتقديم معلومات جديدة لم تكن معروفة مسبقاً في نظم المعلومات التقليدية والتي سميت بالتنقيب في البيانات، فأصبح دورها في هذه البيئة الجديدة ممارسة الدور المعرفي كنظم متكاملة يعمل على توفير معلومات ومعارف واسعة، وتساعد في اتخاذ القرار بشكل أفضل، ويتناول هذا الدور مايلي (Turban et al,2008,p183-186)⁸¹:

1- إيجاد المعرفة وتأمينها (Creating Knowledge): تسعى نظم البحث للمعلومات المحاسبية

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

2- اكتشاف وتصنيف المعرفة (Discovering and Codifying Knowledge): وفرت تقنيات

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

⁷⁹ Michael, J., A. Berry and Gordan S. Linoff, (2000), *Mastering Data Mining*. John Wiley & Sons, In

⁸⁰ Dilnutt, (2002), " KM in practice: three contemporary case studies", *International journal of accounting information systems*, vol.3, issue 2,

⁸¹ Turban, Efraim & Leidner, Dorothy (2008), *Information Technology for Management*, 6th edition, john wiley & sons, NewJersy, P183-186

3- المشاركة بالمعرفة (Sharing Knowledge): فنظم البحث للمعلومات المحاسبية والمصرفية التي توفرها آليات تنقيب التعاون الجماعية تستطيع أن تُساعد العاملين على الوصول والعمل في آن واحد على الوثيقة نفسها، ومن مواقع مختلفة، ومن ثم التنسيق بين نشاطاتهم المختلفة.

4- توزيع المعرفة (Distributing Knowledge): فنظم البحث لتقنيات التنقيب في البيانات وأدوات الاتصال الخاصة بها تستطيع تأمين الوثائق والأشكال الأخرى من المعلومات، وتوزيعها على العاملين في مجال المعلومات والمعرفة، بغرض ربط المكاتب بوحدة الأعمال الأخرى داخل المنظمة وخارجها. ومن ثم فإن الميزات التي توافرها تقنيات التنقيب في البيانات في توفير معلومات دقيقة، وصحيحة، وبشكل سريع، تبين أهمية هذه التقنية في اتخاذ القرار، ومن أهم هذه الميزات على صعيد القرارات الائتمانية في المصارف:

- 1- تساعد متخذ القرار على تجنب منح قروض لعملاء تم التنبؤ بتعثرتهم، واستهداف حملة تسويقية لمقدمي طلبات القروض الجيدة.
- 2- تساعد على تحسين جودة وموثوقية اتخاذ القرارات الاستراتيجية فيما يتعلق بالنشاط الائتماني للبنوك.
- 3- تساعد على زيادة المعرفة، والحد من البدائل، وهذا الأمر يؤدي الى التخلص من حالة عدم التأكد التي تتشكل في هذه البدائل.
- 4- تمكن متخذ القرار في أن يخطط لاستمرارية اجراء الفحص وتحديد المشكلة وعناصرها واتمام الرقابة على ملفات المعلومات والبيانات اللازمة لقراراته.
- 5- تساعد متخذ القرار فيما يحتاج اليه من معلومات لاتخاذ قرارات تسهم في تسريع العمل، وإنجاز المهام، وتبسيط الإجراءات.

الفصل الثالث: الإطار العملي للبحث

يتلخص الهدف من التطبيق في القيام من دراسة حالة (بنك الشام) واستخدام تقنيات التنقيب في البيانات لتحليل بيانات العملاء المستفيدين من تسهيلات ائتمانية للأعوام من 2018م وحتى 2020م باستخدام SQL Server Business Intelligence Development Studio، ومن خلال هذا التحليل يتمكن البنك من فهم خصائص عملائه بشكل أكبر وتمييز العملاء الجيدين والملتزمين بالسداد عن غيرهم فيعدل من أسلوب قراراته عند القيام بعملية المنح الائتماني بما يتوافق مع هذه الخصائص.

أولاً: لمحة عن بنك الشام:

تأسس بنك الشام عام 2006 برأس مال قدره 5 مليار ليرة سورية على شكل شركة مساهمة مغفلة، بموجب قرار الترخيص الصادر عن السيد رئيس مجلس الوزراء في الجمهورية العربية السورية رقم 67/م بتاريخ 2006/09/07.

يعتبر بنك الشام أول مصرف إسلامي في سورية يتخذ الشريعة الإسلامية منهجاً له، وتخضع أنشطة وعمليات البنك لرقابة مصرف سورية المركزي وهيئة الرقابة الشرعية في بنك الشام. يقدم بنك الشام العديد من المنتجات والخدمات ومنها (بنك الشام، 2020)⁸²:

- 1- منتجات الخدمات الالكترونية كخدمات الموبايل البنكي وخدمة البطاقة المصرفية والرسائل النصية.
- 2- منتجات وخدمات الحسابات كفتح حسابات تحت الطلب والحسابات الاستثمارية.
- 3- بالإضافة الى خدمات التمويل كتمويلات الأفراد وتمويلات الشركات.

تتألف الجهة المسؤولة عن منح الائتمان في بنك الشام وفقاً للهيكل التنظيمي للبنك من الجهات الرقابية المتمثلة بإدارة التدقيق الداخلي وإدارة الرقابة والتدقيق الشرعي وإدارة المخاطر وإدارة مراقبة الائتمان والمراقب المصرفي الداخلي وكذلك كافة الجهات التنفيذية المتمثلة بإدارة التمويل التجاري ومركز التمويل التجاري في الإدارة الإقليمية وإدارة التمويل الشخصي والمشاريع الصغيرة وإدارة تنفيذ التمويل، وفروع البنك.

⁸² بنك الشام، البيانات المالية (2020)، .www.chambank.com

ثانياً: الأدوات والبرامج المستخدمة:

1- SQL Server: يعتبر برنامج SQL Server من أهم المزودات لعمل قواعد البيانات من إنتاج شركة Microsoft.

ومن مميزاته يدعم التطوير المتكامل لدورة حياة البيانات في المنظمة (التكامل، التحليل، التقارير) هذه الإمكانية تحدد نتائج التنقيب في البيانات من الاختيار لبيانات محددة وقراءة هذه النتائج والتوقعات بالنسبة للمؤسسة ككل، و يساعد على بناء تطبيقات ذكية من عمليات الإدخال مروراً بالمعالجة والتعديل والحذف الى عرض النتائج والتوقعات، كما أنه يعتبر من البرامج السهل تعامل معه وخاصة لغير المختصين حيث يمكنهم استعراض النتائج وتقييم أداء ودقة الخوارزميات من خلال الخصائص المتواجدة في البرنامج والتي تسهل على الباحثين فهم النموذج المولد.

2- Microsoft SQL Analysis Server (SSAS) : يقوم بتزويد الأدوات للتنقيب في البيانات Data Mining والتي تقوم بتعريف القواعد والنماذج في البيانات.

إنشاء مشاريع خدمات التحليل (Creating an analysis services project) : لإنشاء للتنقيب في البيانات Data Mining نقوم أولاً بإنشاء analysis services project بحيث يتم تشكيل أو اعداد لمصادر البيانات وعرض مصدر البيانات، ومصدر البيانات يقوم بتعريف مصدر الاتصال ومعلومات التحويل التي تتصل بمصدر البيانات الموجودة في نموذج التنقيب في البيانات Data Mining الأساسية كإضافة هياكل Adding mining structures و تنقيب لمشروع تحليل الخدمات Analysis server project، وبعد إضافة Analysis server project يمكن إضافة واحدة أو أكثر من الهياكل التنقيبية.

3- Microsoft Excel 2016 : الملف الذي يحتوي على البيانات الأساسية للعملاء.

ثالثاً: التحليل والخوارزميات المستخدمة:

1- تحليل المشكلة:

تجمع المؤسسات المالية مقداراً كبيراً من البيانات ولكن لا يتم الاستفادة منها، والغرض من هذه البيانات هو الحصول على المعلومات القيمة والوصول الى المعرفة، لذلك فإن التقنية الأحدث والأسرع للتعامل مع البيانات هي التنقيب في البيانات. وكما يعد التنقيب في البيانات طريقة مبتكرة لاكتساب رؤية جديدة وقيمة من خلال تحليل المعلومات الموجودة في قاعدة البيانات. وتتمثل المشكلة في تفادي التعثر لعملاء بنك الشام وذلك من خلال استخراج البيانات من مجموعة بيانات الائتمان للبنك وتحويلها الى هيكل جديد وقيم ومفهوم من خلال تحديد سيناريوهات ائتمانية محفوفة بالمخاطر وتسهيل عملية اتخاذ القرار الائتماني.

سيتم تطبيق التنقيب في البيانات على المعلومات المتعلقة بعملاء الائتمان الخاصة ببنك الشام والتي تم الحصول عليها من النظام المصرفي، ومن المتوقع أن تحدد القواعد التي ستساعد في تصنيف القروض فيما إذا كانت رديئة أم جيدة، وذلك من خلال الكشف عن أنماط الائتمانات الرديئة والجيدة من خلال التنقيب في البيانات.

2- جمع البيانات:

تم الحصول على البيانات من بنك الشام في سوريا في صورة ملف EXCEL. من إدارتي المخاطر والتمويل (البيانات خاصة بعمليات التمويل الشخصي والتجاري من عام 2018 وحتى 2020).

3- عينة البحث:

شملت عينة البحث 1987 سجل لعملاء مستفيدين من تسهيلات تجارية وشخصية.

4- وصف وتحليل البيانات:

- 1- الفرع: ويتضمن الفرع المتواجد ضمن المحافظات والتي تم منح العميل منه ويتضمن (الرئيسي-حلب العزيرية – حلب الفيصل-اللاذقية –صحنايا -المزرعة – الحريقة –حماه-حمص –درعا-المزة – طرطوس).
- 2- تصنيف التمويل: وتتضمن تصنيف نوع التمويل (تجاري – شخصي).
- 3- نوع التمويل: وتتضمن أنواع التمويلات (قرض شخصي – قرض عقاري – قرض سيارة).
- 4- تاريخ المنح: تاريخ منح التمويل.
- 5- تاريخ الاستحقاق: تاريخ استحقاق القرض.
- 6- طريقة السداد: ويوضح كيفية التزام العميل بتسديد قيمة التمويل (شهري – 2شهر – 3أشهر...).
- 7- العملة: والتي توضح نوع عملة التمويل (ليرة سورية- دولار أمريكي- يورو – درهم اماراتي...).
- 8- رهن سيارة: ويوضح في حال تقديم العميل لضمانة سيارة (نعم، لا)
- 9- ضمانة نقدية: ويوضح في حال تقديم العميل لضمانة نقدية (نعم، لا)
- 10- ضمانة عقارية: ويوضح في حال تقديم العميل لرهن عقاري (نعم، لا)
- 11- كفالة شركات: ويوضح في حال العميل مكفول من قبل شركات (نعم، لا)
- 12- كفالة أشخاص: ويوضح فيما إذا كان العميل مكفول من قبل أفراد طبيعيين (نعم، لا)
- 13- تعامل مع بنوك أخرى: ويوضح فيما إذا كان العميل مستفيد من تسهيلات من بنوك أخرى (نعم، لا).
- 14- تاريخ الميلاد: وتتضمن تاريخ ميلاد العميل.
- 15- الجنس: جنس العميل (ذكر، أنثى).
- 16- مهنة العميل: وتتضمن مهنة التي يمارسها العميل.
- 17- عنوان العميل: وتتضمن المحافظة التي يقطن فيها العميل.
- 18- الحالة الاجتماعية للعميل: وتتضمن الوضع الاجتماعي للعميل (متزوج – أعزب – مطلق – أرمل).
- 19- مستوى التعليم: وتتضمن مستوى التعليم (دكتوراه – ماجستير – جامعي – ثانوي-اعدادي – ابتدائي).
- 20- تصنيف العميل: وتتضمن تصنيف العميل فيما إذا كان متعثر أم غير متعثر.

5- تهيئة البيانات:

قواعد بيانات العالم الحقيقي معرضة بشكل كبير للبيانات الصاخبة والمفقودة وغير المتسقة نظراً لحجمها الضخم وتعدد مصادرها الغير متجانسة. لذلك فإن المعالجة المسبقة للبيانات هي خطوة مهمة وشرطية في عملية التنقيب في البيانات، وكما يؤدي تحليل البيانات التي لم يتم فحصها بعناية الى نتائج مضللة وغير

صحيحة (Kamber et al,2006)⁸³، وتواجد الكثير من البيانات غير ذات الصلة والمتكررة أو البيانات الصاخبة وغير الموثوقة ستكون أكثر صعوبة في اكتشاف المعرفة أثناء مرحلة التدريب. يمكن أن تستغرق خطوات إعداد البيانات وتصفيتهما قدرًا كبيراً من وقت المعالجة ولكنها مرحلة حتمية ولا يمكن تخطيها. وتشمل الأنشطة خلال هذه المرحلة تنظيف البيانات، اختيار السمة، تحويل البيانات وتجميعها وتنسيقها (Pyle,1999)⁸⁴.

1.5.تنظيف البيانات (Data Cleaning):

تعمل إجراءات تنظيف البيانات على ملء القيم المفقودة، وتحييد القيم المتطرفة أو ازلتها، وحل التناقضات، وكما تلعب جودة البيانات دوراً مهماً في المنظمات الموجهة نحو المعلومات. الاتساق والاكتمال والدقة والصلاحية والتوقيت هي من الخصائص المهمة لجودة البيانات لذلك من المهم الحصول على بيانات جيدة لاستخراج المعرفة (Feysia,2018)⁸⁵.

وقد تم اجراء عملية تنظيف البيانات على العينة التي تم الحصول عليها من النظام المعلوماتي لبنك الشام لاحتوائها على بعض القيم المتكررة والقيم الصاخبة، تم الإجراء على النحو التالي:

- اجراء توحيد على الحقول التي تحتوي على عدة ادخالات كتوحيد على الحقول التالية: (الفرع – نوع التمويل – مهنة العميل – مستوى التعليم – الجنس).
- تم حذف السجلات المتكررة.
- تم التحقق من عدم وجود تعارضات بين الحقول كمستوى التعليم والمهنة.

2.5.تحويل البيانات (Data Transformation):

بمجرد تجميع البيانات وإصلاح مشاكل البيانات الرئيسية فلا بد من تحويل البيانات للتحليل، وذلك يتضمن إضافة بعض الحقول المشتقة واجراء عمليات تجميع وتطبيع وتقدير وبناء السمات (Kamber et al,2006)⁸⁶.

⁸³ Han, J. and Kamber, M. (2006). *Data Mining: Concepts and Techniques. Second Edition, Morgan Kaufmann Publishers, San Francisco*

⁸⁴ Pyle, D., 1999 *Data Preparation for Data Mining Morgan Kaufmann Publisher.*

⁸⁵ Feysia, Ketema (2018) " The use of data mining to predict the loan repayment risk: the case of Oromia credit and saving share company ", **A thesis**, college of natural science, Addis Ababa University.

⁸⁶ Han, J. and Kamber, M. (2006). *Data Mining: Concepts and Techniques. Second Edition, Morgan Kaufmann Publishers, San Francisco*

ومن أجل جعل مجموعة البيانات المستخدمة لهذه الدراسة مناسبة لعملية التنقيب في البيانات تم استخدام مجموعة من طرق تحويل البيانات، ومنها:

- تم تحويل عناوين الحقول وبيانات الحقول الموضوع بالعربي الى اللغة الإنكليزية.
- تم اشتقاق عمر العميل من تاريخ الميلاد.
- تم اشتقاق مدة التمويل من تاريخ المنح وتاريخ الاستحقاق.

3.5.3 تمثيل قيمة السمة واشتقاقها (Attribute value representation and derivation):

تم اجراء عملية تمثيل لبعض حقول البيانات كإعطائها رموز أو أرقام، وكانت على الشكل التالي:

● الحقول المتعلقة بخصائص التمويل:

| Value representaion | IsOtherBank | Value representaion | Stage | Value representaion | Period_of_loan | Value representaion | Period_of_payment | Value representaion | BranchID |
|---------------------|-------------|---------------------|-------|---------------------|--------------------------------|---------------------|-------------------|---------------------|--------------|
| 0 | نعم | NotDefault | 1 | 1 | أقل أو تساوي ٢ سنوات | 1-M | 1-MONTHS | 1 | اللاذقية |
| 1 | لا | Default | 3 | 2 | أكثر من ٢ أقل أو تساوي ٥ سنوات | 2-M | 2-MONTHS | 2 | الرئيسي |
| | | | | 3 | أكثر من ٥ سنوات | 3-M | 3-MONTHS | 3 | المرزة |
| | | | | | | 5-M | 5-MONTHS | 4 | حلب الفيصل |
| | | | | | | 6-M | 6-MONTHS | 5 | حماة |
| | | | | | | | | 6 | درعا |
| | | | | | | | | 7 | حمص |
| | | | | | | | | 8 | صحنايا |
| | | | | | | | | 9 | طرطوس |
| | | | | | | | | 10 | المرزعة |
| | | | | | | | | 11 | الحرثية |
| | | | | | | | | 12 | فرع العزيزية |

جدول (1) تمثيل سمات خصائص التمويل

● الحقول المتعلقة بالضمانات:

| Value representation | Personal Guarantie | Value representation | Mortgage Guarantie | Value representation | Equipment Guarantie | Value representation | Corporate Guarantie | Value representation | Cash Collateral | Value representation | Cargarantie |
|----------------------|--------------------|----------------------|--------------------|----------------------|---------------------|----------------------|---------------------|----------------------|-----------------|----------------------|-------------|
| 1 | نعم | 1 | نعم | 1 | نعم | 1 | نعم | 1 | نعم | 1 | نعم |
| 0 | لا | 0 | لا | 0 | لا | 0 | لا | 0 | لا | 0 | لا |

جدول (2) تمثيل سمات خصائص الضمانات

• الحقول المتعلقة بالعميل:

| Value representation | Gender | Value representation | AgentAge | Value representation | Civil status |
|----------------------|--------|----------------------|----------|----------------------|--------------|
| M | ذكر | 1 | 25-0 | 1 | MARRIED |
| F | أنثى | 2 | 50-26 | 2 | SINGLE |
| | | 3 | 75-51 | 3 | DIVORCED |
| | | 4 | وأكثر 76 | 4 | WIDOWED |

جدول (3) تمثيل سمات خصائص العميل

4.5. اختيار السمات (Attribute selection):

إن اختيار السمات يقلل من حجم مجموعة البيانات عن طريق إزالة أو استبعاد السمات غير ذات الصلة والزائدة عن الحاجة، كما أن التنقيب في البيانات في مجموعة منخفضة من السمات له فائدة إضافية حيث أنه يقلل من عدد السمات التي تظهر الأنماط المكتشفة، مما يساعد على تسهيل فهم الأنماط⁸⁷.

ويتم اختيار سمات التنبؤ الخاصة بالتنبؤ بسداد القروض وتعثرها من منظور الأعمال التجارية للمؤسسة وأهداف البحث⁸⁸، والسمات التي تم اعتمادها في البحث كالتالي:

⁸⁷ Han, J. and Kamber, M. (2006). *Data Mining: Concepts and Techniques. Second Edition*, Morgan Kaufmann Publishers, San Francisco

⁸⁸ Witten H. Ian and Frank, E. (2005). *Data Mining: Practical Machine Learning Tools and Techniques, Second Edition*. Morgan Kaufmann Publishers, San Francisco

| No | Attribute | Data type | Data type in Database | Attribute selection |
|----|-------------------|-----------|-----------------------|---------------------|
| 1 | ID | Numeric | int | Selected |
| 2 | ACCOUNT_CLASS | String | nvarchar | Selected |
| 3 | ACCOUNT_TYPE | String | nvarchar | Selected |
| 4 | PROFESSION | String | nvarchar | Selected |
| 5 | period_of_payment | String | nvarchar | Selected |
| 6 | BranchId | Numeric | nvarchar | Not selected |
| 7 | stage | String | nvarchar | Selected |
| 8 | IsOtherBank | Numeric | bit | Selected |
| 9 | CarGarantie | Numeric | bit | Selected |
| 10 | CashCollateral | Numeric | bit | Selected |
| 11 | CorporateGarantie | Numeric | bit | Selected |
| 12 | EquipmentGarantie | Numeric | bit | Not selected |
| 13 | LIMITAMOUNT | Numeric | float | Not selected |
| 14 | MortgageGarantie | Numeric | bit | Selected |
| 15 | PersonalGarantie | Numeric | bit | Selected |
| 16 | AgentActivity | String | nvarchar | Not selected |
| 17 | Gender | String | varchar | Selected |
| 18 | Address | String | nvarchar | Not selected |
| 19 | PeriodOfLoan | Numeric | nvarchar | Selected |
| 20 | AgnetAge | Numeric | nvarchar | Selected |

*int: لتخزين الأعداد الصحيحة

*nvarchar: هو معيار حرفي بطول غير محدد ويمثل بلغات متعددة (العربية- الإنكليزية - اليونانية...)

*varchar: هو معيار حرفي بطول غير محدد وهو مناسب مع اللغة الإنكليزية

*bit: لتخزين قيم منطقية إما نعم أو لا

*float: لتخزين الأعداد بوجود الفاصلة العشرية

جدول (4) تمثيل اختيار السمات

تم تحديد 15 سمة لتدريبها وهي: رمز التعريف، تصنيف التمويل، نوع التمويل، المهنة، طريقة السداد، تصنيف العميل، تعاملات مع بنوك أخرى، ضمانات سيارة، ضمانات نقدية، كفالة شركات، ضمانات عقارية، كفالة شخصية، الجنس، مدة التمويل، عمر العميل.

6- الخوارزميات المستخدمة:

التصنيف (Classification) :

يستخدم التصنيف بشكل واسع في حل الكثير من المشكلات خاصة تلك التي تتعلق بالأعمال من خلال تحليل مجموعة من البيانات ووضعها على شكل أصناف أو أقسام يمكن استخدامها فيما بعد لتصنيف البيانات مستقبلاً، وهنا يكمن الفرق بين التصنيف (Classification) وإنشاء العناقيد (Clustering)، فالتصنيف

يقصد به تقسيم البيانات الى مجاميع يتم تحديدها مسبقاً، أما إنشاء العناقيد (Clustering) فهو يعني تقسيم البيانات الى مجاميع ليست معروفة مسبقاً.

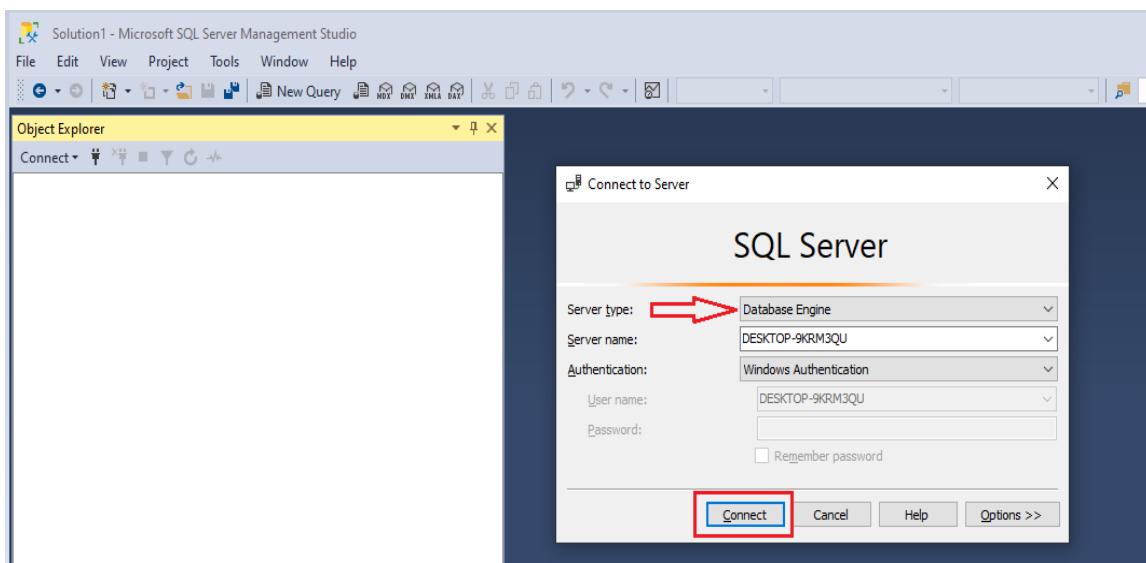
فهناك عدد من الطرق التي يمكن استخدامها في تصنيف البيانات باستخدام خوارزميات مختلفة مثل الخوارزميات الإحصائية (Statistical Algorithms) والشبكات العصبية (Neural Network) و الخوارزميات الجينية (Genetic Algorithms) وطريقة الجار الأقرب (Nearest Neighbor method) و نظرية بيبز المبسط (Naïve Bayes) و طريقة شجرة القرار (Decision Trees) وهي هيكل شجري يقدم مجموعة من القرارات التي تولد قواعد لمجموعة البيانات المصنفة (data classification) وكما تشترط هذه التقنية على وجود حقل قرار يتم تصنيف البيانات بناء عليه.

والخوارزميات التي سيتم اعتمادها في هذا البحث هي طرق شجرة القرار (Decision Trees) ونظرية بيبز المبسط (Bayes Naive).

رابعاً: مرحلة التنفيذ:

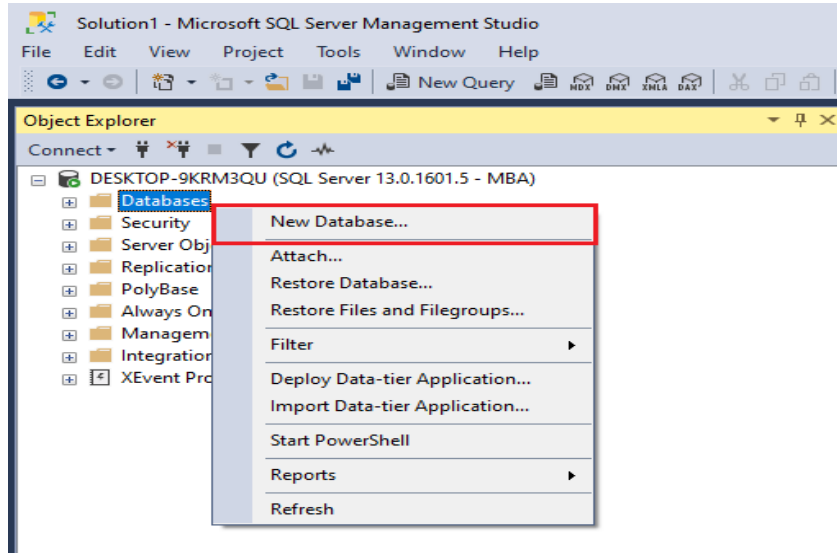
1- مرحلة بناء قاعدة البيانات:

يتوافر العديد من بيئات قواعد البيانات التي تم تصميمها من قبل شركات عالمية متخصصة لتوفير أفضل الوسائل والأدوات والأساليب اللازمة لإدارة قواعد البيانات، تم الاعتماد في هذا البحث على برنامج MS SQL SERVER لبناء قواعد البيانات لاستخدامها لاحقاً في عملية التنقيب في البيانات، وكانت بناءً على الخطوات التالية:



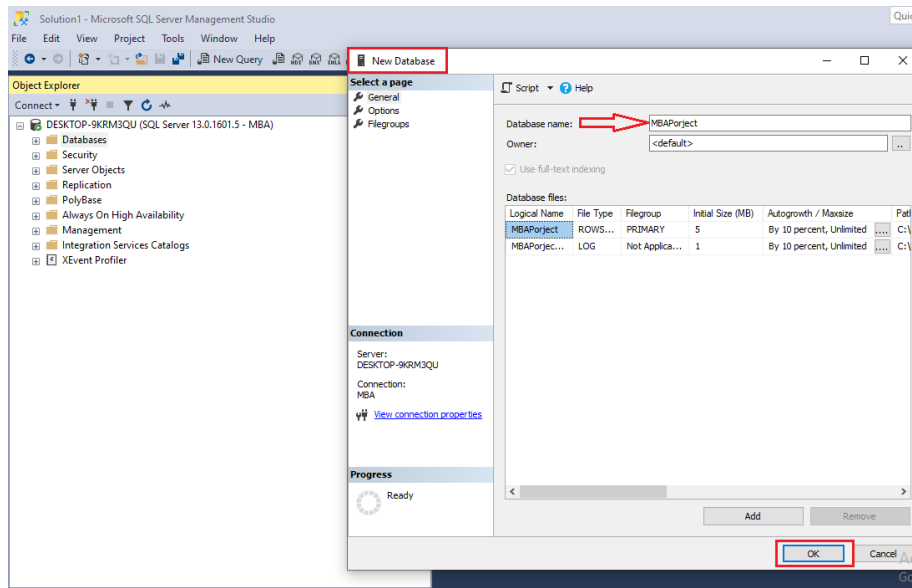
شكل (3) واجهة البرنامج المستخدم SQL SERVER

عند Server type يتم وضع Database Engine والمقصود به النوع المراد هو قاعدة بيانات، و Server name يتم وضع مكان تواجد قواعد البيانات وفي حالة البحث هو اسم الحاسوب، و Authentication يتم اختيار Windows Authentication ليتم المصادقة على المستخدم الذي تم الدخول من خلاله الى الحاسوب (Windows user)، ومن ثم نضغط على زر Connect .



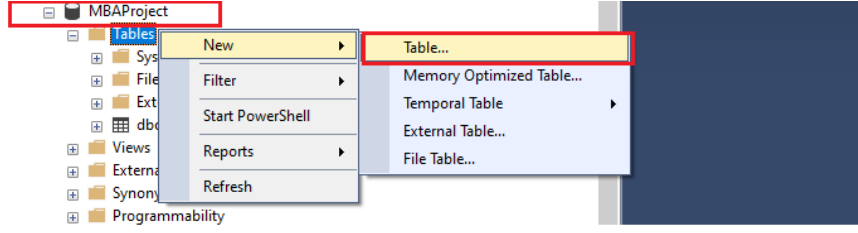
شكل (4) انشاء ملف جديد لقاعدة البيانات

بعد تحديد الخيارات السابقة تظهر لنا الواجهة الموجودة بالشكل (4) وعند خيار Database نضغط خيار .New Database



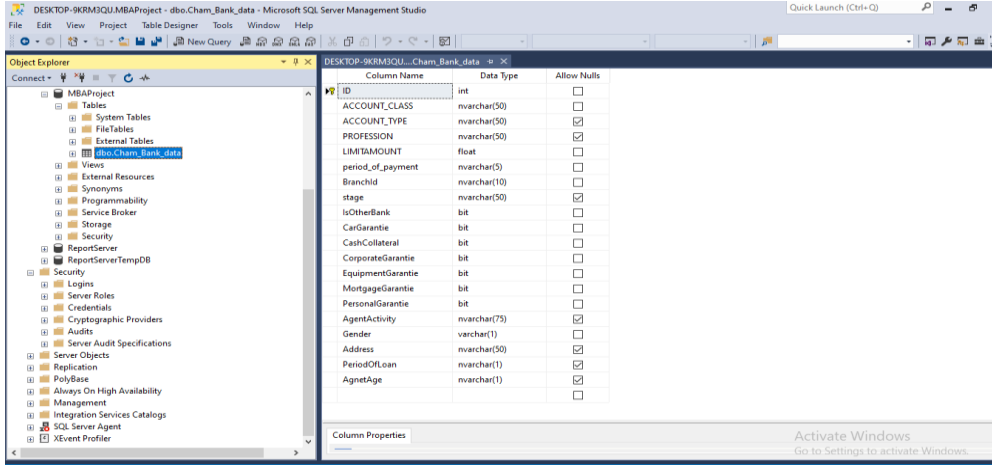
شكل (5) واجهة خيارات انشاء ملف لقواعد البيانات

يتم تسمية قاعدة البيانات الجديدة عند خيار Database name ومن ثم نضغط على زر ok .



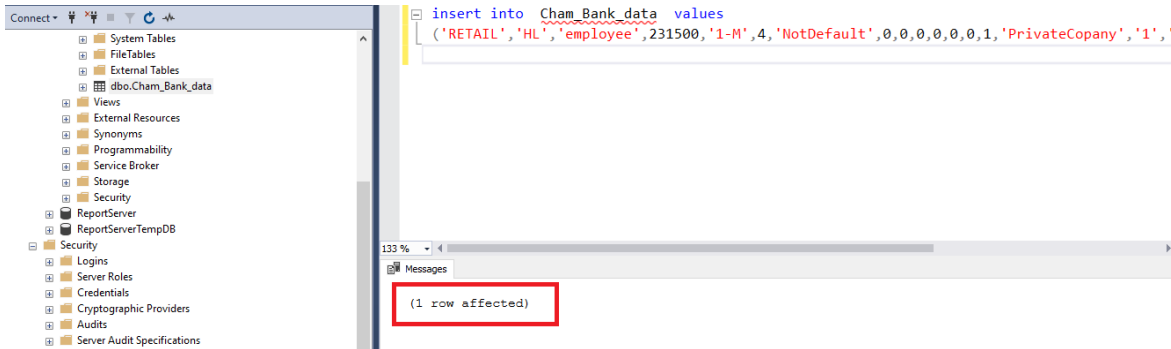
شكل (6) انشاء جدول ضمن قاعدة البيانات

بعد انشاء قاعدة بيانات جديدة يتم انشاء جدول لتخزين البيانات فيها وتتم وفقاً كم هو مبين في الشكل (6).



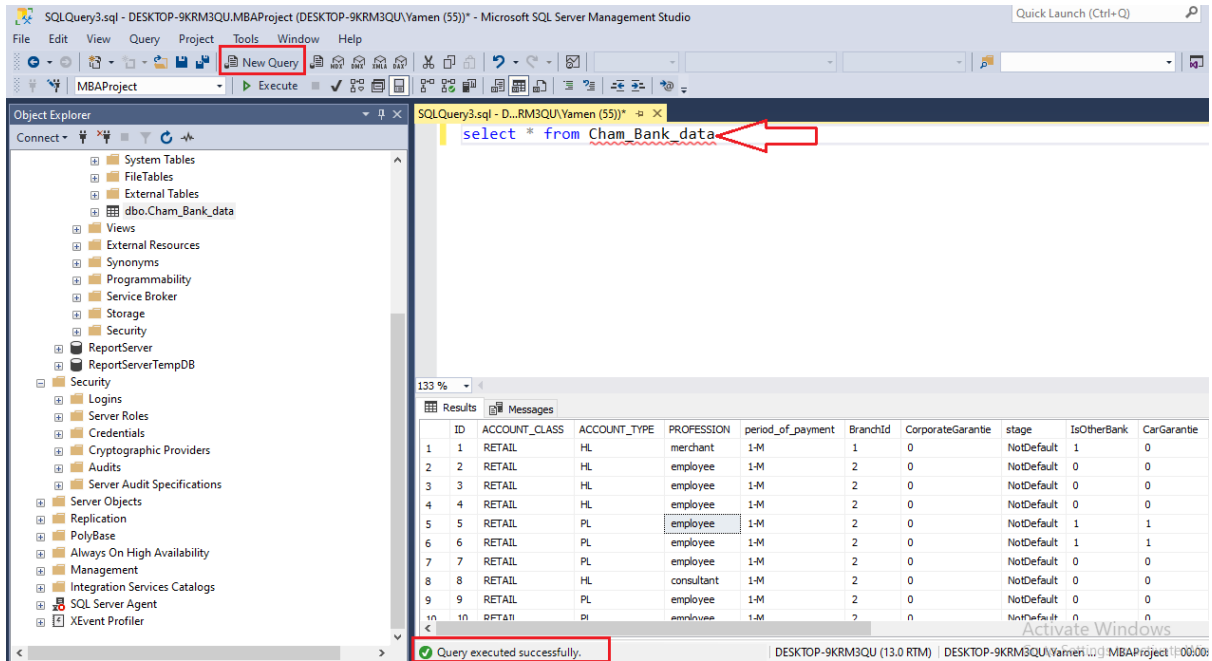
شكل (7) تعريف وتحديد خصائص الحقول ضمن الجدول

تمت تسمية الجدول Cham_Bank_data ومن ثم تم تعريف عناوين الحقول وتمثيل السمات الخاصة بكل حقل كما تم شرحها في الجدول رقم (4).



شكل (8) ادخال البيانات الى الجدول

يتم إدراج البيانات إلى الجدول وذلك من البيانات التي تمت تهيئتها مسبقاً كما هو موضح بالشكل (8)، ونستخدم هنا تعليمة SQL لإضافة سطر جديد الى الجدول الذي تم إنشاؤه ويتم تكرار العملية على كافة السجلات المهيئة في عملية تنظيف البيانات⁸⁹.



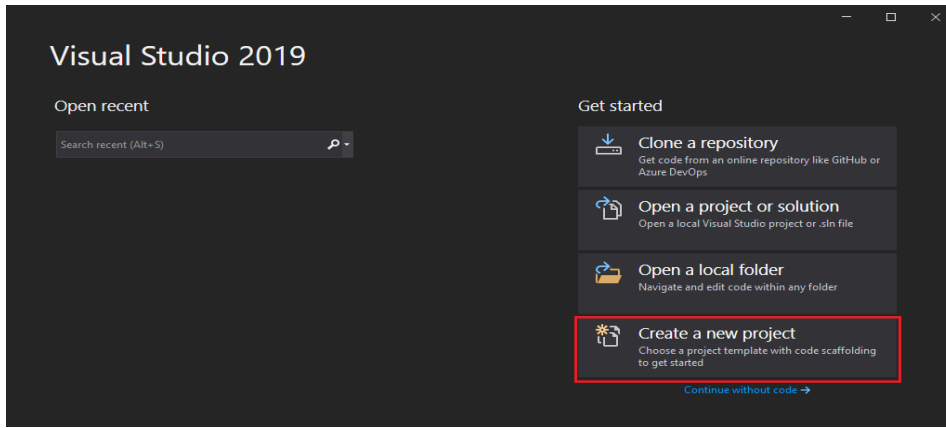
شكل (9) انشاء استعلام للبيانات

باستخدام SQL QUERY نستعلم عن البيانات المدخلة في المرحلة السابقة وذلك للتحقق من عملية ادخال كافة البيانات المطلوبة بالشكل الصحيح.

2- تهيئة بيئة العمل لاختيار الخوارزميات:

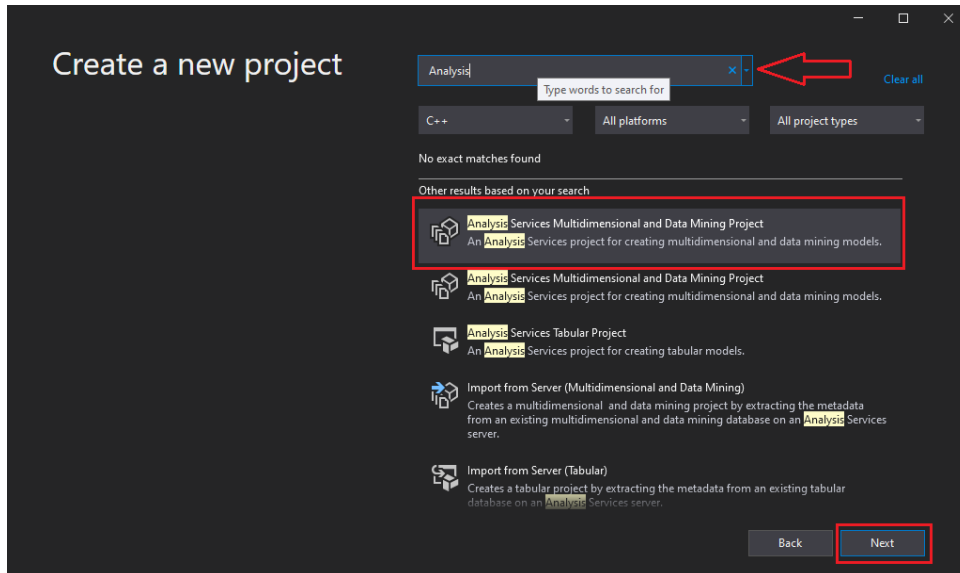
هناك العديد من البرامج يمكن استخدامها للتعامل مع قواعد البيانات والتي تساعد في عملية التنقيب في البيانات تم استخدام برنامج Visual Studio 2019 وهو من منتجات شركة Microsoft.

⁸⁹ <https://www.w3schools.com/sql/default.asp>



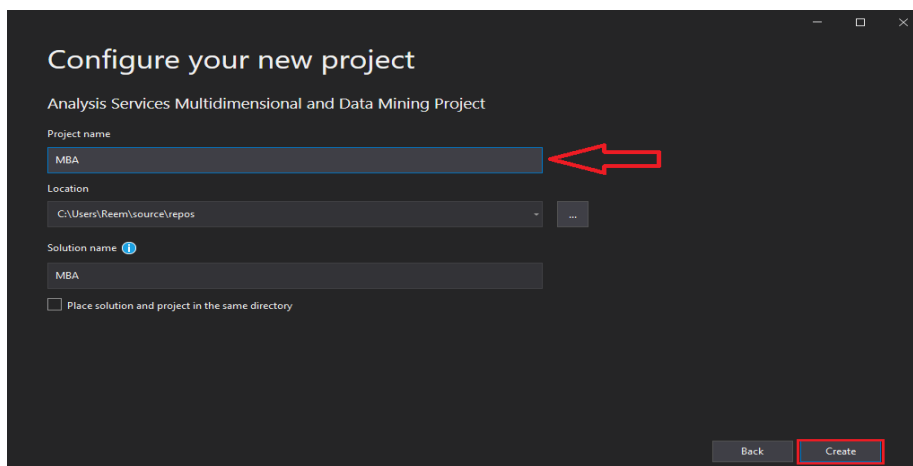
شكل (10) واجهة برنامج Visual Studio 2019

نقوم باختيار Create a new project لخلق مشروع جديد المراد التنقيب في بياناته.



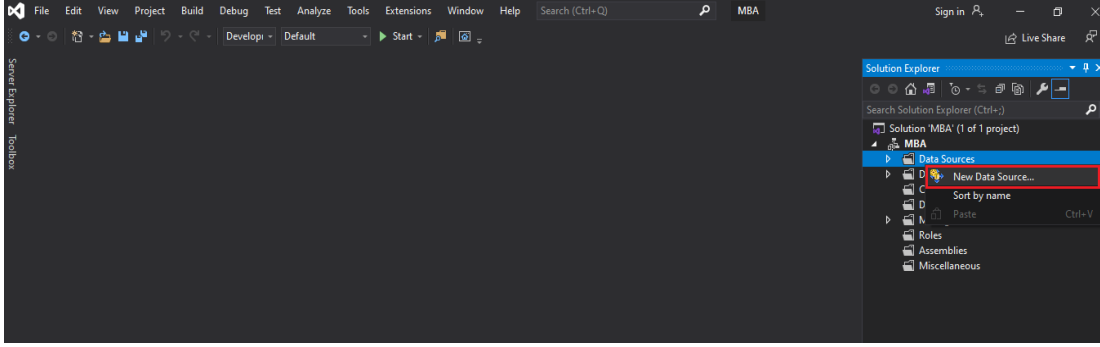
شكل (11) واجهة اختيار نموذج التنقيب في البيانات

نقوم بالبحث عن نموذج Analysis services multidimensional and Data mining project ومن ثم ضغط Next.



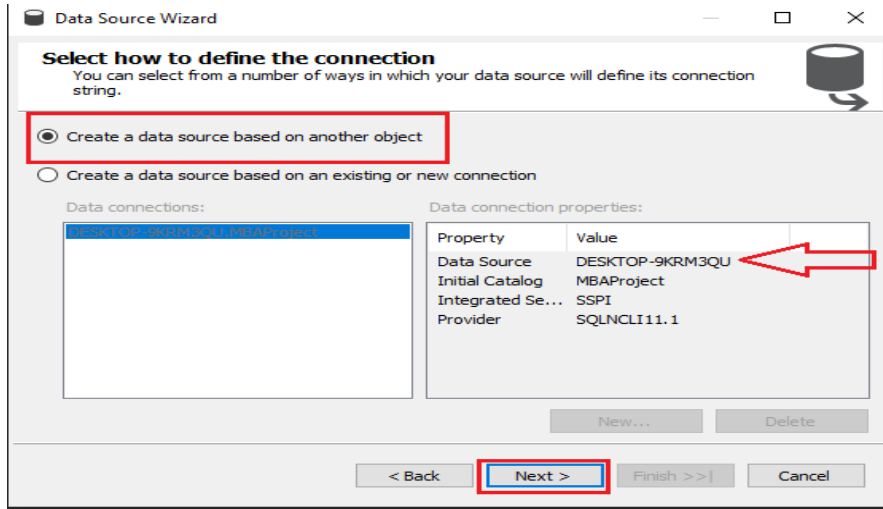
شكل (12) واجهة تسمية المشروع

نقوم بتسمية المشروع وفي حالة هذا البحث تمت تسميته MBA ومن ثم نضغط Create .

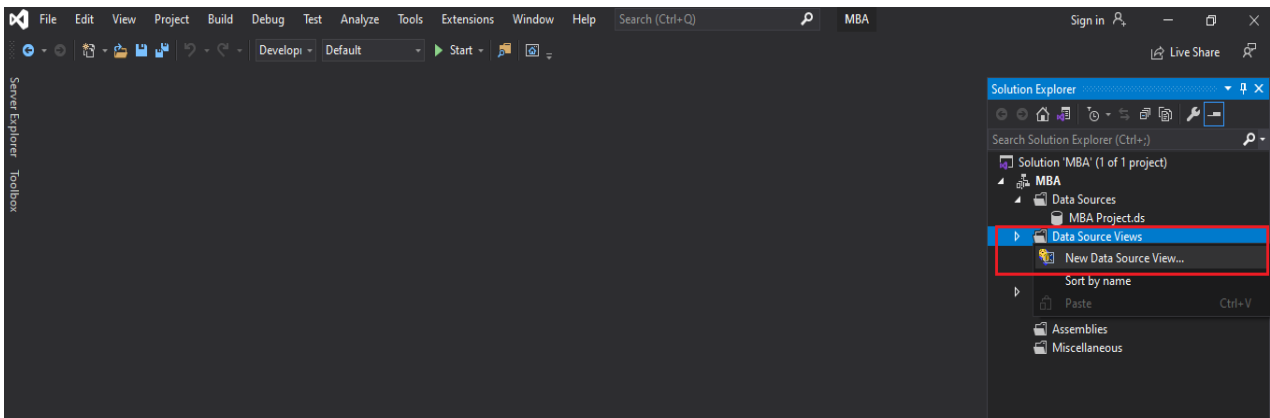


شكل (13) واجهة New Data Source

بعد انشاء خلق مشروع جديد سنكون بحاجة للوصل مع قاعدة البيانات، نقوم بالضغط بالزر اليميني على حقل Data Sources ومن ثم نضغط على New Data Source كما هو مبين في الشكل (13).

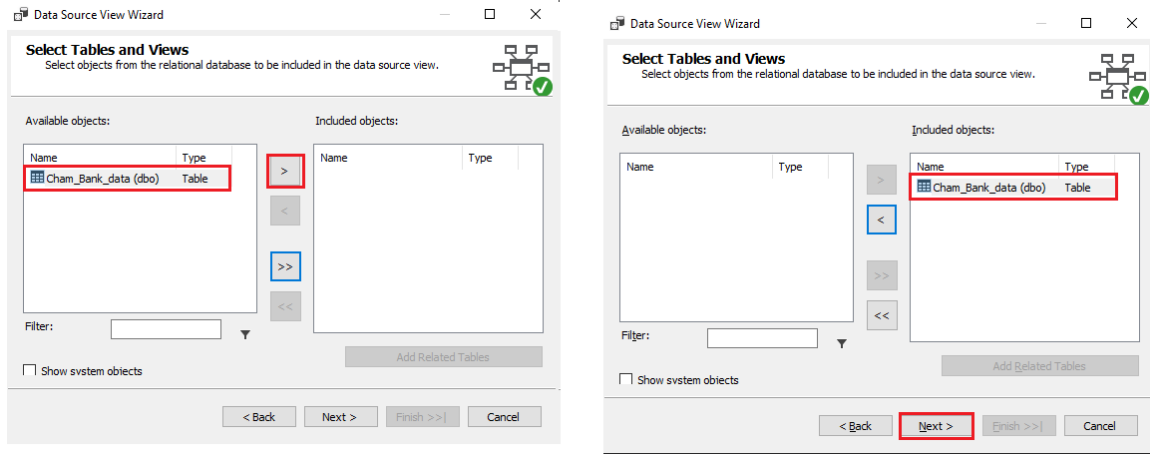


شكل (14) آلية الوصل مع قاعدة البيانات



شكل (15) واجهة Data Source views

بعد الانتهاء من مرحلة الوصل مع قاعدة البيانات نقوم بتحديد الجدول ضمن قاعدة البيانات والذي يحتوي على البيانات المراد التنقيب فيها.



شكل (16) اختيار الجدول من قاعدة البيانات

ولإنهاء مرحلة تهيئة بيئة العمل لاختيار الخوارزميات نقوم بتحديد الجدول ضمن قاعدة البيانات كما هو موضح بالشكل (16) وذلك يكون من ضمن خيارات Data Source views كما هو موضح بالشكل (15).

3-مرحلة تطبيق خوارزمية شجرة القرار:

شجرة القرار هي نموذج استكشافي يظهر على شكل شجرة، ويمثل كل فرع من فروعها سؤالاً تصنيفياً وتمثل أوراقها أجزاء من قاعدة البيانات تنتمي للتصنيفات التي تم بناؤها. إن العقدة في الشجرة هي عبارة عن فحص لبعض الصفات، أما الفرع فهو قيمة محتملة للصفة ويكون التصنيف حسب الخطوات التالية⁹⁰:

1. ابدأ من الجذر.
2. افحص الصفة.
3. تحرك الى أسفل فرع الشجرة.

-مراحل بناء خوارزمية شجرة القرار:

يتم بناء شجرة القرار بمرحلتين:

1- التعليم: وهي خطوة بناء نموذج التصنيف، ويتم في هذه المرحلة بتدريب خوارزمية التصنيف (classifying algorithm) على بيانات التدريب (training data) المحتوية على سجلات معروفة لبناء المصنف الذي يستخدم لفحص بيانات الفحص (data test) التي تحتوي على سجلات غير معروفة.

⁹⁰ المبارك، البدوي سعد البدوي.(2017): "استخدام تقنيات التنقيب في البيانات لاستكشاف أنماط مؤثرات التحصيل الأكاديمي لطلاب المرحلة الثانوية"رسالة ماجستير غير منشورة، كلية الدراسات العليا، جامعة النيلين.

2- **التصنيف:** وفي هذه الخطوة يتم استخدام النموذج ليتنبأ بفئات البيانات غير المعروفة، ويتم في هذه المرحلة تقييم أداء المصنف بحساب عدد السجلات المتوقعة المصنفة بشكل صحيح والسجلات المصنفة بشكل خطأ فيما يسمى بمصفوفة التعارض (confusion matrix) ويتم تقييم أداء نماذج المصنفات بالحصول على خوارزميات تصنيف تسعى للحصول على أعلى دقة وأقل نسبة خطأ عند تطبيقها على بيانات الاختبار⁹¹.

- حساب مصفوفة التعارض (confusion matrix) :

تتم جدولة عدد السجلات المصنفة بشكل صحيح وعدد السجلات المصنفة بشكل خاطئ على شكل مصفوفة تعرف بمصفوفة التعارض أو التضارب.

| ACTUAL CLASS | PREDICTED CLASS | | |
|--------------|-----------------|-----------|-----------|
| | Class = 1 | Class = 1 | Class = 0 |
| Class = 1 | TP | FN | |
| Class = 0 | FP | TN | |

جدول (5) يوضح مصفوفة التعارض

مدخل F_{ij} في مصفوفة التعارض يشير الى عدد السجلات في class i المتوقع أن يكون في class j مثلاً f_{01} يشير الى عدد السجلات في class 0 التي تم توقعها بشكل خطأ في class 1.

اعتماداً على المصفوفة فإن: مجموع السجلات المتوقعة بشكل صحيح هي $(f_{11}+f_{00})$

مجموع السجلات المتوقعة بشكل خطأ هي $(f_{01}+f_{10})$

وتحسب درجة الدقة Accuracy لتقييم أداء المصنفات اعتماداً على مصفوفة التعارض من المعادلة التالية:

$$Accuracy = \frac{\text{Number of correct predictions}}{\text{Total number of predictions}} = \frac{f_{11} + f_{00}}{f_{11} + f_{10} + f_{01} + f_{00}}$$

- مميزات استخدام خوارزمية شجرة القرار:

- 1- تعتبر خوارزمية شجرة القرار طريقة سهلة لا تحتاج الى وسطاء لبناء نماذج التصنيف أي أنها لا تحتاج الى فرضيات سابقة فيما يتعلق بنوع التوزيعات الاحتمالية.
- 2- إن التقنيات المطورة من أجل بناء شجرة القرار غير مكلفة، مما يعين إمكانية بناء نماذج سرعة حتى عندما يكون حجم مجموعة التدريب كبيرة جداً.

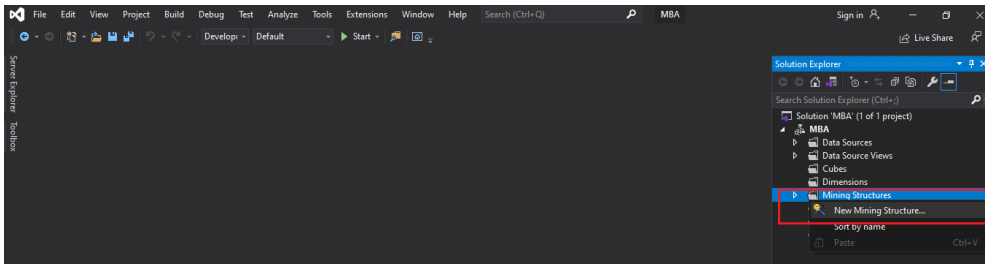
⁹¹ أبو القاسم أحمد، هبة. (2015): "دراسة أثر تطبيق تقنية التصنيف على قطاع التأمين (دراسة حالة شركة شيكان)" رسالة ماجستير غير منشورة، كلية الدراسات العليا، جامعة النيلين.

- 3- من السهل نسبياً تفسير أشجار القرار إذا كانت صغيرة الحجم كما يمكن مقارنة دقة الأشجار مع تقنيات أخرى من أجل العديد من مجموعة البيانات.
- 4- تعتبر أشجار القرار قوية وحتى بوجود تشويبه (Noise).

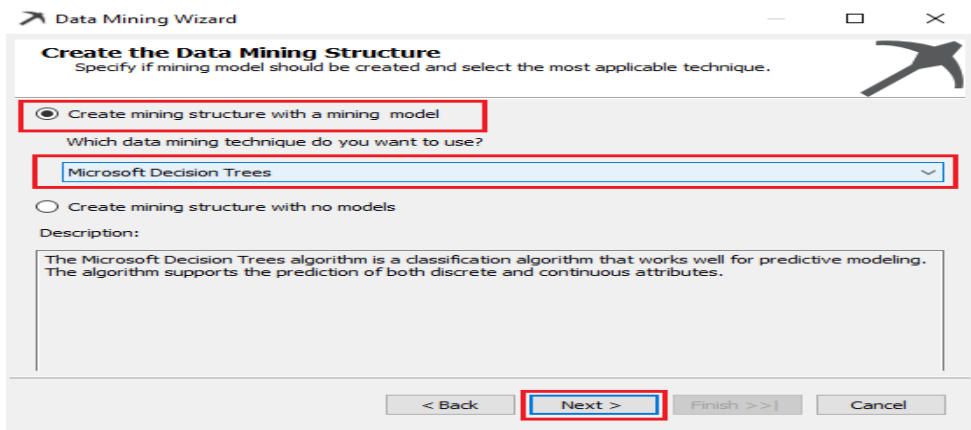
- تطبيق شجرة القرار على برنامج Visual Studio 2019:

سيتم تطبيق خوارزمية شجرة القرار لأول مرة على بيانات عام 2018 والذي يحتوي على 1,224 سجل لعملاء حاصلين على تسهيلات، ومن ثم سيتم مقارنة النموذج مع العملاء الحاصلين على تسهيلات جديدة عام 2019 لنبين فيه دقة النموذج عند مقارنته مع بيانات واقعية، ومن ثم سيتم بناء نموذج ثاني يحتوي على بيانات عام 2018 و 2019 ومقارنته مع عام 2020، ومن ثم سيتم بناء نموذج نهائي يحتوي على بيانات 2018 و2019 و2020.

لبناء نموذج التنقيب نختار من Mining structure، Microsoft Decision trees كما هو موضح في الشكل (17) و(18).

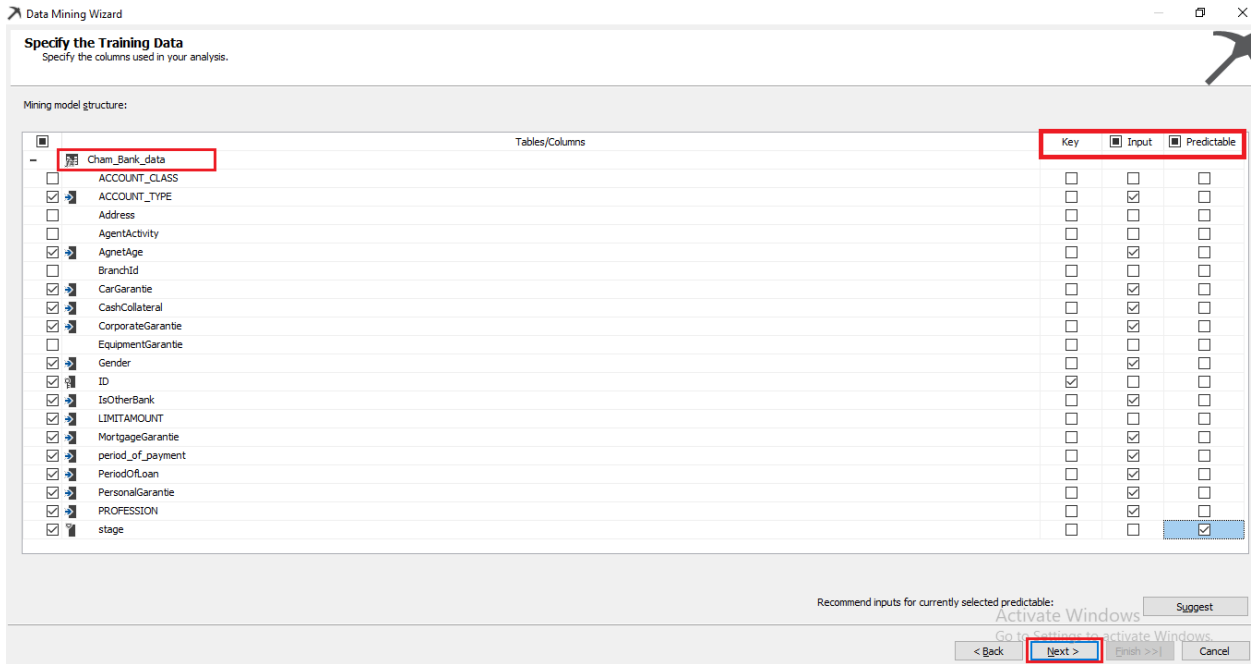


شكل (17) بناء نموذج التنقيب



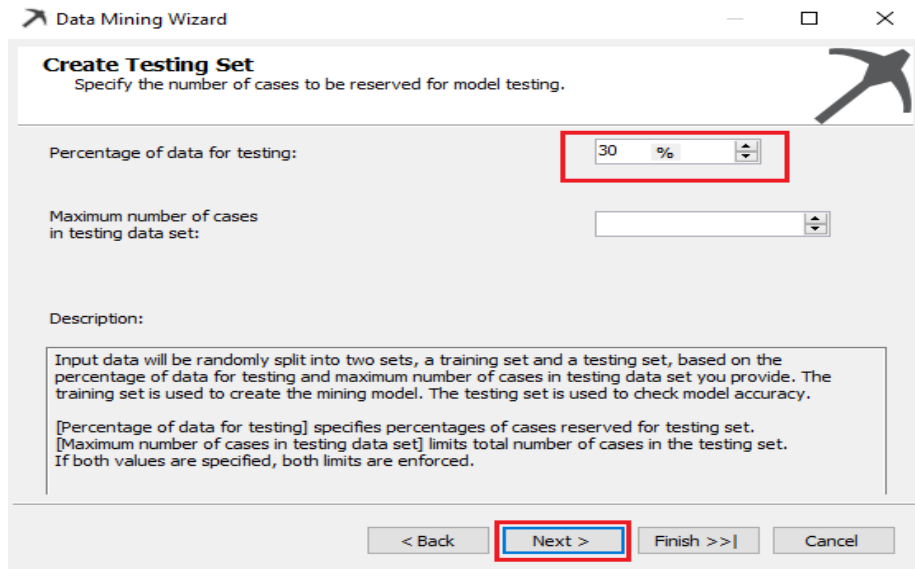
شكل (18) اختيار خوارزمية شجرة القرار

ومن ثم نقوم بتحديد الواصفات المراد تحديدها كإدخال (input) للخوارزمية والواصفات المراد استبعادها والواصفات المراد التنبؤ بها كما هو موضح بالشكل (19).



شكل (19) تحديد الوصفات

نقوم بتحديد نسبة البيانات المراد اختبارها بعد الانتهاء من إعداد النموذج وقد تم تحديد نسبة 30% وهي القيمة الافتراضية للبرنامج وتعتبر نسبة 50% هي الحد الأعلى للبيانات التي يمكن اختبارها، وكما تساعد نسبة بيانات الاختبار في تحديد درجة دقة النموذج (عبيد، 2020).⁹²

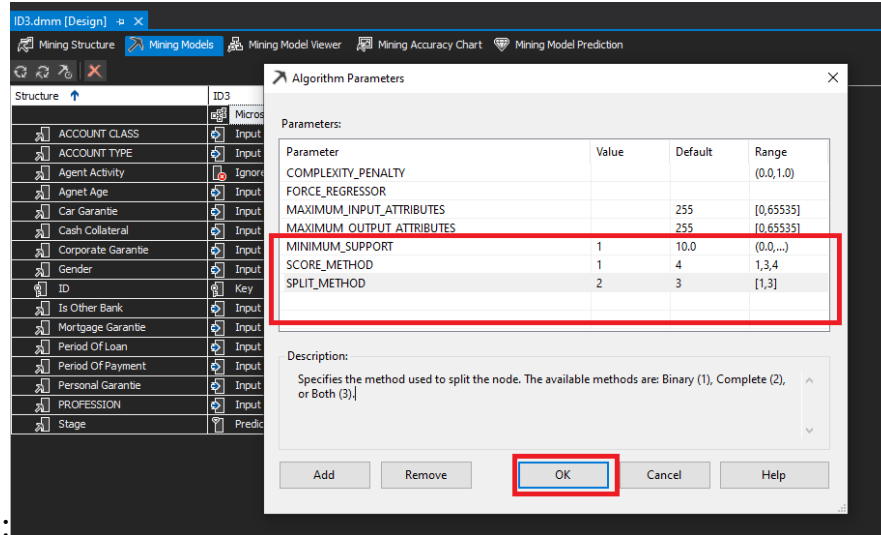


شكل (20) تحديد نسبة البيانات المراد اختبارها

⁹² عبيد، مصطفى. 2020، ورقة عمل بعنوان: التنقيب في قواعد البيانات واستكشاف المعلومات المخبأة فيها — وصف خوارزمياتها وتطبيقاتها في المجتمع.

يتم ضبط متغيرات خوارزمية Microsoft Decision Trees كما في الشكل (21) ويهمنها ما يلي:

- MINIMUM_SUPPORT = 1
- SCORE_METHOD = 1
- SPLIT_METHOD = 2



شكل (21) ضبط متغيرات خوارزمية Microsoft Decision Trees

المتغير MINIMUM_SUPPORT يعني عدد الحالات الواجب تواجدها في أي عقدة في الشجرة وهنا تم وضع القيمة 1 نظراً لأن قاعدة البيانات صغيرة نسبياً⁹³.

المتغير SCORE_METHOD : للاختيار بين ثلاث خوارزميات لتحديد متى يتم فصل عقدة في شجرة القرار إلى عقدتين أو أكثر وهنا كنا قد اخترنا entropy وهي عينة من أمثلة التدريب والتي تحسب بالمعادلة التالية (MacLennan, Tang, & Crivat, 2009):

$$\text{Entropy}(S) = -p_+ \log_2 p_+ - p_- \log_2 p_-$$

- Entropy (1)
- Bayesian with K2 Prior (3)
- Bayesian Dirichlet Equivalent with Uniform Prior (4)

المتغير SPLIT_METHOD يحدد هذا المتغير كيف يتم تقسيم العقدة في الشجرة (MacLennan, Tang, & Crivat, 2009).

- Binary (1) : يتم تقسيم العقدة إلى عقدتين حصراً، بحيث إذا كان لدينا سمة لها ثلاث قيم لسمة

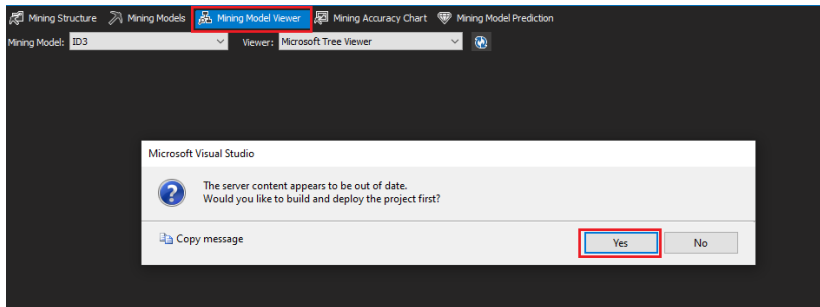
MacLennan, J., Tang, Z., & Crivat, B. (2009). *Data Mining with Microsoft SQL Server 2008*. Indianapolis,, Indiana, ⁹³ United States of America: Wiley Publishing, Inc.

Exam Amali=good, Exam Amali=not تصبح Exam-Amali (good, average, poor)

good

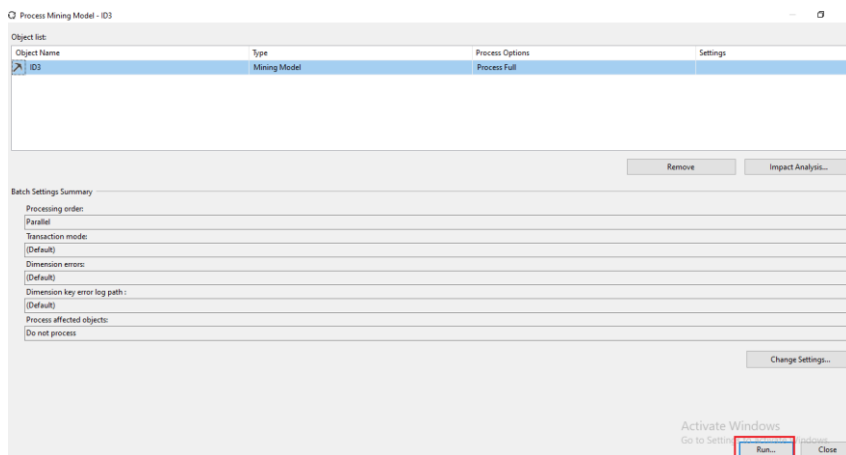
- Complete (2) : أي يتم انقسام العقد لكل القيم الممكنة لها فالسمة التي تملك قيمتين تنقسم لفرعين والتي تملك ثلاث قيم تنقسم لثلاث وهكذا.
- Both (3) : أي الخيارين السابقين معاً والخوارزمية هي التي تحدد أيهما أفضل بشكل أوتوماتيكي.

ومن ثم ننتقل الى نافذة Mining Model viewer والتي تطلب رسالة تأكيد على بناء النموذج كما هو موضح بالشكل (22).



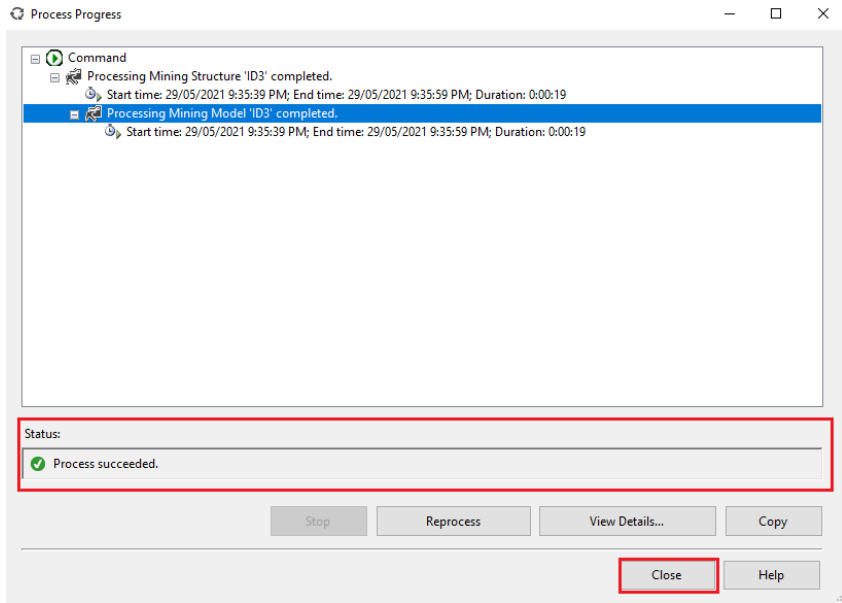
شكل (22) تأكيد بناء النموذج

وبعد تأكيد بناء النموذج تظهر نافذة لإكمال عملية بناء النموذج وعكسه على Analysis server وعند الضغط على زر Run يبدأ ببناء النموذج.



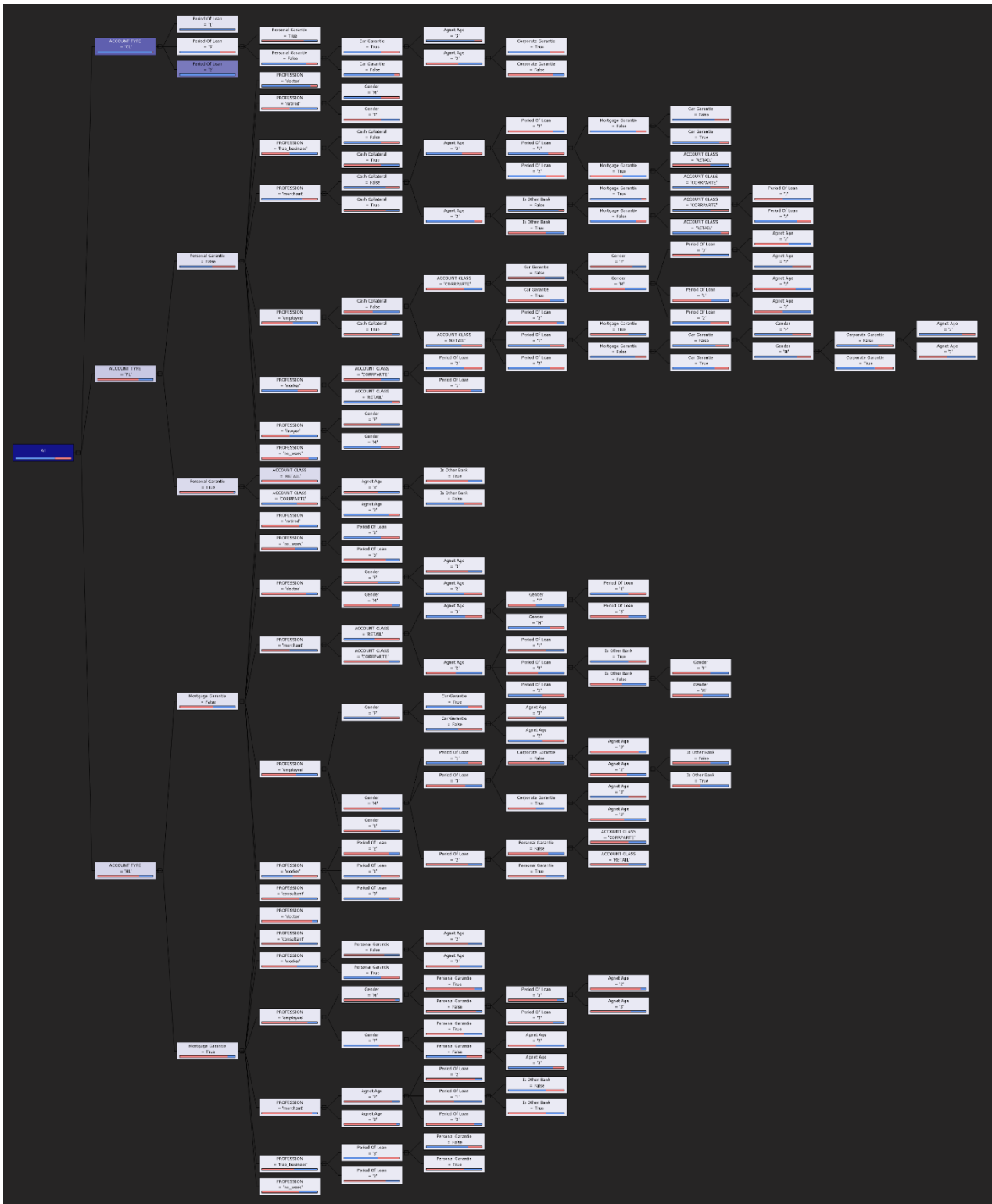
شكل (23) اكمال بناء النموذج

ومن ثم تظهر نافذة نرى من خلالها عملية البناء والتنفيذ وعند الانتهاء سيظهر رسالة بإتمام العملية بنجاح في حال عدم وجود أخطاء، وفي حال وجود أي خطأ سيظهر نوع الخطأ المسبب بعدم إكمال العملية، وكما تظهر هذه النافذة مدة تنفيذ العملية كما هو موضح بالشكل (24).



شكل (24) تنفيذ النموذج

ومن ثم سيظهر لنا الشكل البياني لشجرة القرار كما هو مبين في شكل (25)



شكل (25) شجرة القرار عن بيانات 2018

يمثل الشكل (25) المعرفة والقواعد المستنبطة من البيانات المدخلة وتتمثل بشكل شجرة بحيث تكون الأوراق فيها هي النتيجة النهائية لدعم قرار ما، والذي يساعد في معرفة حالات العملاء الجدد بناء على بيانات لعملاء سابقين، أي في حال قام العميل بطلب تسهيل جديد سيتم تمرير خصائصه على هذه الشجرة وسيتم التنبؤ

بحالته هل سيتعثر أم لا والذي سيساعد في اتخاذ قرار منح ائتماني سليم، وذلك من خلال مقارنته مع عملاء حاصلين على تسهيلات سابقة وحاملين على نفس الخصائص.

يتم قراءة القواعد من شجرة القرار على شكل if-then بعض الأمثلة على القواعد التي تمت استنتاجها من النموذج الأول لشجرة القرار:

- If (account type "CL" and Period of loan "1") then "Default"
- If (account type "CL" and Period of loan "3" and Personal Guarantie "true") then "Default"
- If (account type "PL" and Personal Guarantie "true" and Account Class "Retail") then "NotDefault"

| Predicted | NotDefault (Actual) | Default (Actual) |
|------------|---------------------|------------------|
| NotDefault | 88 | 15 |
| Default | 21 | 243 |

جدول (6) مصفوفة التعارض لشجرة القرار عن بيانات 2018

يبين الجدول (6) مصفوفة التعارض الذي تقوم بجدولة السجلات المصنفة بشكل صحيح مع السجلات المصنفة بشكل خاطئ، وذلك من خلال أخذ عينة من البيانات واختبارها مع النموذج ومقارنه التوقعات مع الحالة الواقعية، وعند بناء النموذج تم تحديد النسبة المراد اختبارها والموضحة بالشكل (20)، فقد تم اختبار 30% من البيانات المدخلة عن عام 2018 والتي تمثل 367 سجل من أصل 1224 سجل.

وكما تبين عدد السجلات التي كانت في الواقع غير متعثرين وتم التوقع بعدم تعثرهم 88 سجل ما يعادل 23.9% من العينة، وعدد السجلات في الواقع هم غير متعثرين وتم التنبؤ بتعثرهم 21 سجل أي ما يعادل 5.7%، وعدد السجلات التي في الواقع متعثرين وتم التنبؤ بعدم تعثرهم 15 سجل ما يعادل 4.1%، بينما تم التوقع بتعثر 243 سجل وهم في الواقع متعثرين أيضاً أي ما يعادل 66.21%، ومنه نستنتج بأن نسبة دقة النموذج بعد اختبار جزء من العينة المدخلة 90.2%.

| rateGara... | Gender | IsOtherBank | MortgageGarantie | period_of_pay... | PeriodOfLoan | PersonalGarantie | PROFESSION | stage | Prediction |
|-------------|--------|-------------|------------------|------------------|--------------|------------------|------------|------------|------------|
| | F | False | False | 1-M | 1 | True | employee | NotDefault | NotDefault |
| | M | False | False | 1-M | 1 | True | employee | NotDefault | NotDefault |
| | M | False | False | 1-M | 1 | True | employee | NotDefault | NotDefault |
| | M | False | False | 1-M | 1 | True | employee | NotDefault | NotDefault |
| | M | False | False | 1-M | 1 | True | employee | NotDefault | NotDefault |
| | F | False | False | 1-M | 1 | True | employee | NotDefault | NotDefault |
| | M | False | False | 1-M | 1 | True | employee | NotDefault | NotDefault |
| | M | False | False | 1-M | 1 | True | employee | NotDefault | NotDefault |
| | M | True | False | 1-M | 1 | True | employee | NotDefault | NotDefault |
| | M | False | False | 1-M | 1 | True | employee | NotDefault | NotDefault |
| | M | False | False | 1-M | 1 | True | employee | NotDefault | NotDefault |
| | M | False | False | 1-M | 1 | True | employee | NotDefault | NotDefault |
| | F | False | False | 1-M | 1 | True | employee | NotDefault | NotDefault |
| | F | False | False | 1-M | 1 | True | employee | NotDefault | NotDefault |
| | M | False | False | 1-M | 1 | True | employee | NotDefault | NotDefault |
| | M | False | False | 1-M | 1 | True | employee | NotDefault | NotDefault |
| | F | False | False | 1-M | 1 | True | employee | NotDefault | NotDefault |
| | F | False | False | 1-M | 1 | True | employee | NotDefault | NotDefault |
| | M | False | False | 1-M | 1 | True | employee | NotDefault | NotDefault |
| | M | False | False | 1-M | 1 | True | employee | NotDefault | NotDefault |
| | F | False | False | 1-M | 1 | True | employee | NotDefault | NotDefault |
| | M | False | False | 1-M | 1 | True | employee | NotDefault | NotDefault |
| | M | False | False | 1-M | 1 | True | employee | NotDefault | NotDefault |
| | F | False | False | 1-M | 1 | True | employee | NotDefault | NotDefault |
| | M | False | False | 1-M | 1 | False | worker | NotDefault | Default |

شكل (26) مقارنة العملاء الجدد 2019 مع قواعد شجرة القرار

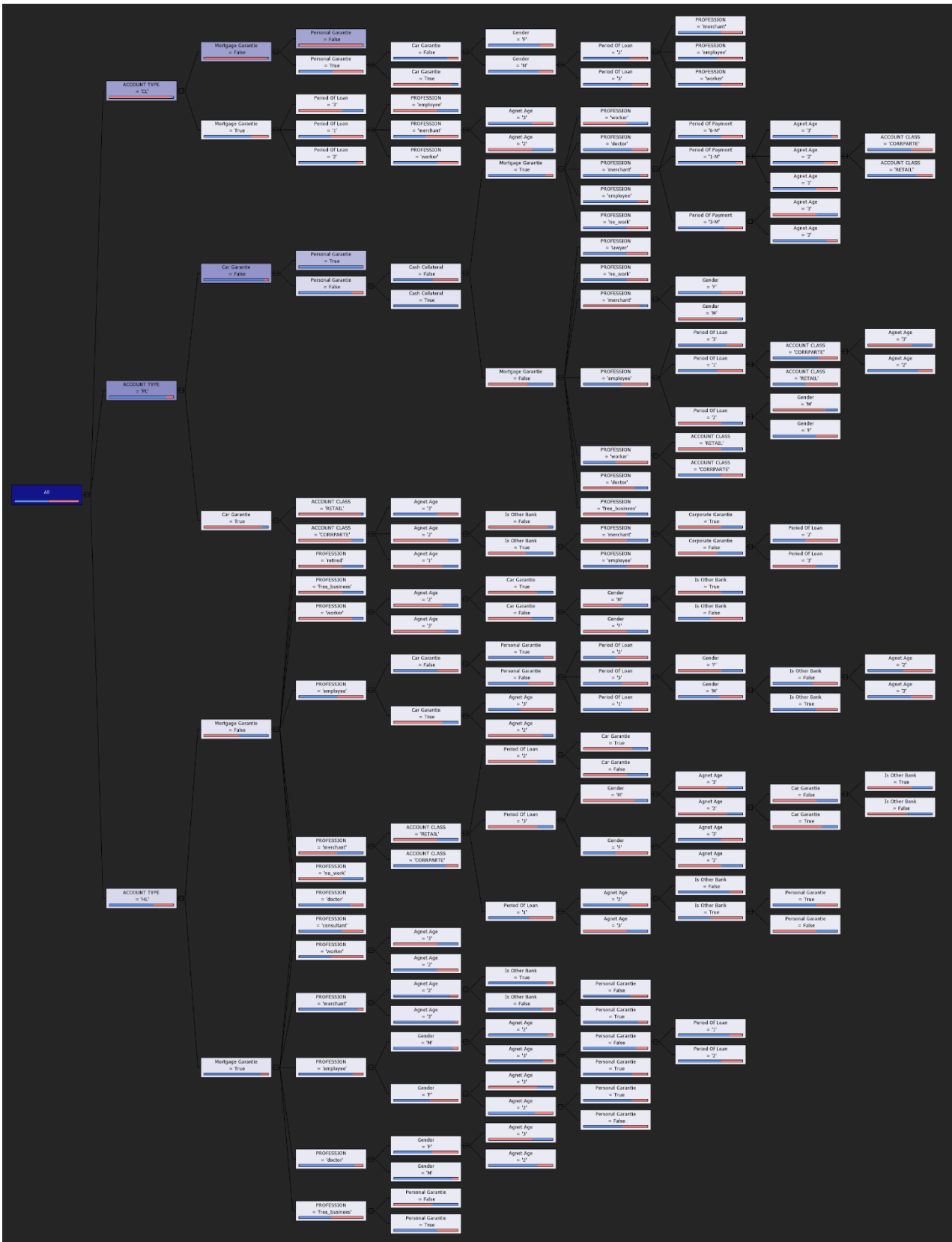
تم تطبيق قواعد شجرة القرار على العملاء الحاصلين على تسهيلات جديدة خلال عام 2019 كما هو وارد في الشكل (26) وكانت النتائج وفق ما هو وارد في الجدول (7):

| predicted | Default(Actual) | NotDefault(Actual) |
|------------|-----------------|--------------------|
| Default | 0 | 31 |
| NotDefault | 3 | 491 |

جدول (7) نتائج مقارنة العملاء الجدد 2019 مع شجرة القرار

ومنه نستنتج بأن درجة دقة النموذج مقارنة مع العملاء الجدد 2019 هي 93.52% أي تم التنبؤ ب 491 سجل من أصل 524 بينما لم يتمكن النموذج بالتنبؤ بشكل صحيح بنسبة 6.48%.

بعد الانتهاء من بناء النموذج الأول لشجرة القرار بناء على بيانات 2018 واستنباط القواعد وتطبيقها على بيانات 2019 واستنتاج درجة دقة النموذج، سيتم تطبيق النموذج الثاني من شجرة القرار بناء على بيانات 2018 و 2019 بشكل كامل واستخراج القواعد وتطبيقها على بيانات 2020 لاستنتاج درجة دقة النموذج، وسيتم تطبيق نفس خطوات بناء النموذج الأول.



شكل (27) شجرة القرار عن بيانات 2018 و2019

| Predicted | NotDefault (Actual) | Default (Actual) |
|------------|---------------------|------------------|
| NotDefault | 244 | 17 |
| Default | 22 | 250 |

جدول (8) مصفوفة التعارض لشجرة القرار عن بيانات 2018 و2019

يبين الجدول (8) مصفوفة التعارض وقد تم اختبار 30% من البيانات المدخلة عن عامي 2018 و2019 والتي تمثل 533 سجل من أصل 1779 سجل.

وكما تبين عدد السجلات التي كانت في الواقع غير متعثرين وتم التوقع بعدم تعثرهم 244 سجل ما يعادل 45.8% من العينة، وعدد السجلات في الواقع هم غير متعثرين وتم التنبؤ بتعثرهم 22 سجل أي ما يعادل 4.1%، وعدد السجلات التي في الواقع متعثرين وتم التنبؤ بعدم تعثرهم 17 سجل ما يعادل 3.2%، بينما تم التوقع بتعثر 250 سجل وهم في الواقع متعثرين أيضاً أي ما يعادل 46.9%، ومنه نستنتج بأن نسبة دقة النموذج بعد اختبار جزء من العينة المدخلة 92.7%.

| ACCOUNT | Age | Age | CarGarant | CashColla | Corporate | Gender | ID | IsOtherBa | Mortgage | period_of | PeriodOfI | PersonalC | PROFESSI | stage | PRED | RESULT |
|---------|-----|-------|-----------|-----------|-----------|--------|-------|-----------|----------|-----------|-----------|-----------|------------|------------|-------|--------|
| PL | 2 | TRUE | FALSE | FALSE | M | 1 | TRUE | FALSE | 1-M | 3 | FALSE | employee | NotDefault | Default | FALSE | |
| HL | 3 | FALSE | FALSE | FALSE | M | 2 | FALSE | TRUE | 1-M | 3 | FALSE | merchant | NotDefault | NotDefault | TRUE | |
| PL | 2 | FALSE | FALSE | FALSE | M | 3 | FALSE | FALSE | 1-M | 3 | TRUE | employee | NotDefault | NotDefault | TRUE | |
| PL | 2 | TRUE | FALSE | FALSE | M | 4 | TRUE | FALSE | 1-M | 3 | TRUE | employee | NotDefault | Default | FALSE | |
| PL | 2 | TRUE | FALSE | FALSE | M | 5 | FALSE | FALSE | 1-M | 3 | TRUE | employee | NotDefault | Default | FALSE | |
| HL | 2 | FALSE | FALSE | FALSE | M | 6 | FALSE | TRUE | 1-M | 3 | FALSE | employee | NotDefault | NotDefault | TRUE | |
| HL | 2 | FALSE | FALSE | FALSE | M | 7 | FALSE | FALSE | 1-M | 3 | FALSE | employee | NotDefault | NotDefault | TRUE | |
| PL | 2 | FALSE | FALSE | FALSE | M | 8 | FALSE | FALSE | 1-M | 3 | TRUE | employee | NotDefault | NotDefault | TRUE | |
| PL | 2 | FALSE | FALSE | FALSE | M | 9 | FALSE | FALSE | 1-M | 3 | TRUE | employee | NotDefault | NotDefault | TRUE | |
| PL | 2 | FALSE | FALSE | FALSE | F | 10 | FALSE | FALSE | 1-M | 3 | TRUE | employee | NotDefault | NotDefault | TRUE | |
| PL | 2 | FALSE | FALSE | FALSE | M | 11 | FALSE | TRUE | 1-M | 3 | FALSE | employee | NotDefault | NotDefault | TRUE | |
| PL | 2 | FALSE | FALSE | FALSE | F | 12 | FALSE | FALSE | 1-M | 3 | TRUE | employee | NotDefault | NotDefault | TRUE | |
| PL | 2 | FALSE | FALSE | FALSE | M | 13 | FALSE | TRUE | 1-M | 3 | TRUE | employee | NotDefault | NotDefault | TRUE | |
| PL | 2 | FALSE | FALSE | FALSE | F | 14 | FALSE | FALSE | 1-M | 3 | TRUE | employee | NotDefault | NotDefault | TRUE | |
| PL | 2 | FALSE | FALSE | FALSE | M | 15 | FALSE | TRUE | 1-M | 3 | TRUE | employee | NotDefault | NotDefault | TRUE | |
| PL | 2 | FALSE | FALSE | FALSE | M | 16 | FALSE | FALSE | 1-M | 3 | TRUE | employee | NotDefault | NotDefault | TRUE | |
| PL | 2 | FALSE | FALSE | FALSE | M | 17 | FALSE | FALSE | 1-M | 3 | TRUE | employee | NotDefault | NotDefault | TRUE | |
| PL | 2 | TRUE | FALSE | FALSE | M | 18 | FALSE | FALSE | 1-M | 3 | FALSE | employee | NotDefault | Default | FALSE | |
| PL | 2 | FALSE | FALSE | FALSE | M | 19 | FALSE | FALSE | 1-M | 3 | TRUE | employee | NotDefault | NotDefault | TRUE | |
| HL | 2 | FALSE | FALSE | FALSE | M | 20 | FALSE | FALSE | 1-M | 3 | FALSE | employee | NotDefault | NotDefault | TRUE | |
| PL | 2 | FALSE | FALSE | FALSE | M | 21 | FALSE | FALSE | 1-M | 3 | TRUE | employee | NotDefault | NotDefault | TRUE | |

شكل (28) مقارنة العملاء الجدد 2020 مع قواعد شجرة القرار

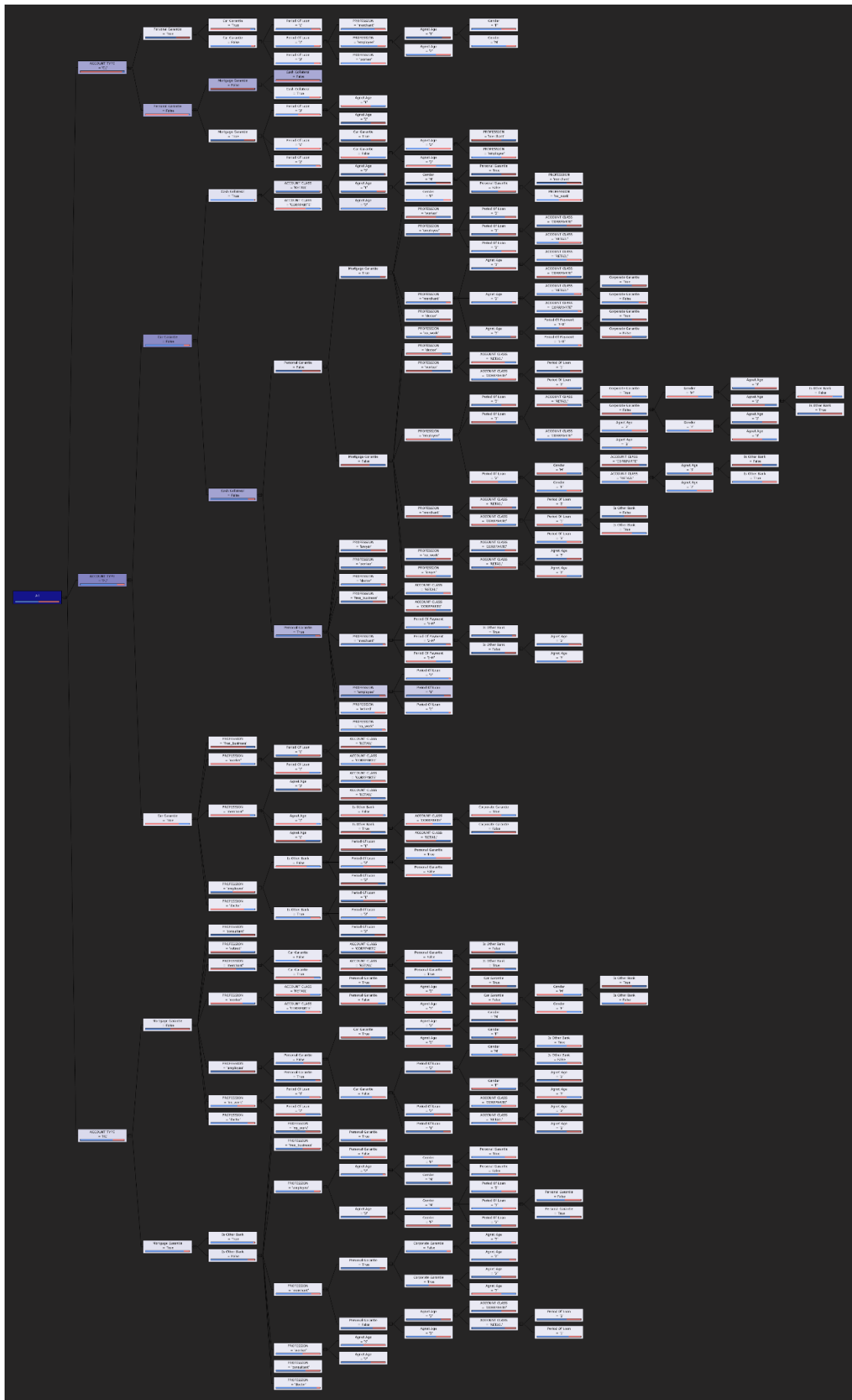
تم تطبيق قواعد شجرة القرار على العملاء الحاصلين على تسهيلات جديدة خلال عام 2020 كما هو وارد في الشكل (27) وكانت النتائج وفق ما هو وارد في الجدول (9):

| predicted | Default(Actual) | NotDefault(Actual) |
|------------|-----------------|--------------------|
| Default | 4 | 11 |
| NotDefault | 12 | 181 |

جدول (9) نتائج مقارنة العملاء الجدد 2020 مع شجرة القرار

ومنه نستنتج بأن درجة دقة النموذج مقارنة مع العملاء الجدد 2020 هي 88.9% أي تم التنبؤ ب 185 سجل من أصل 208 بينما لم يتمكن النموذج بالتنبؤ بشكل صحيح بنسبة 11.1%.

بعد الانتهاء من بناء النموذج الثاني لشجرة القرار بناء على بيانات 2018 و 2019 واستنباط القواعد وتطبيقها على بيانات 2020 واستنتاج درجة دقة النموذج، سيتم تطبيق النموذج النهائي من شجرة القرار بناء على بيانات 2018 و 2019 و 2020 بشكل كامل واستخراج القواعد، وسيتم تطبيق نفس خطوات بناء النموذج الأول والثاني.



شكل (29) شجرة القرار عن بيانات 2018 و 2019 و 2020

| Predicted | NotDefault (Actual) | Default (Actual) |
|------------|---------------------|------------------|
| NotDefault | 108 | 11 |
| Default | 5 | 74 |

جدول (10) مصفوفة التعارض لشجرة القرار عن بيانات 2018 و 2019 و 2020

يبين الجدول (10) مصفوفة التعارض وقد تم اختبار 10% من البيانات المدخلة عن الأعوام 2018 و 2019 و 2020 والتي تمثل 198 سجل من أصل 1980 سجل، وقد تم اختبار 10% من البيانات بدلاً من 30% كما هو وارد بالنموذج الأول والثاني وذلك لأننا بحاجة الى تدريب أكبر قدر ممكن من البيانات والتحقق من درجة دقة النموذج من خلال 10% من البيانات المتبقية.

وكما تبين عدد السجلات التي كانت في الواقع غير متعثرين وتم التوقع بعدم تعثرهم 108 سجل ما يعادل 54.5% من العينة، وعدد السجلات في الواقع هم غير متعثرين وتم التنبؤ بتعثرهم 5 سجل أي ما يعادل 2.5%، وعدد السجلات التي في الواقع متعثرين وتم التنبؤ بعدم تعثرهم 11 سجل ما يعادل 5.6%، بينما تم التوقع بتعثر 74 سجل وهم في الواقع متعثرين أيضاً أي ما يعادل 37.4%، ومنه نستنتج بأن نسبة دقة النموذج بعد اختبار جزء من العينة المدخلة 91.9%.

4-مرحلة تطبيق خوارزمية بيز المبسط (Naïve Bayes):

تعد خوارزمية تصنيف Naïve Bayes إحدى تقنيات التصنيف التي لا تستخدم أي قاعدة مثل قاعدة شجرة القرار، ويعتبر أساس Naïve Bayes هو نظرية الاحتمالات، وتمثل الطريقة المباشرة لحساب الاحتمال في البحث عن الحدث المتكرر وتصنيف الحالة غير المرئية إلى الحدث المتكرر الحدوث.

وباستخدام أنواع احتمالية أكثر تعقيداً يعطي توقع أفضل للحدث غير المرئي، وللحصول على الاحتمال المسبق نقسم تكرار الحدث الأكثر تكراراً على إجمالي عدد المثيلات. يُطلق على احتمال وقوع حدث ما إذا علمنا أن السمة لها قيمة معينة (أو أن العديد من المتغيرات لها قيم معينة) تسمى الاحتمال الشرطي، ويسمى هذا أيضاً الاحتمال اللاحق لأنه يحسب الاحتمال بعد أن يكتسب المعلومات بينما يحسب الاحتمال المسبق قبل أن تحصل على المعلومات (Zhang, 2004)⁹⁴.

⁹⁴ Harry Zhang "The Optimality of Naive Bayes". FLAIRS2004 conference. (available online: PDF (<http://www.cs.unb.ca/profs/hzhang/publications/FLAIRS04ZhangH.pdf>))

ويعرف الاحتمال الشرطي بأنه احتمال وقوع A من خلال B يساوي الاحتمال المشترك لحدوث كل من A و B ، مقسوماً على احتمال B، وفق المعادلة التالية:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

وباستخدام نظرية بييز المبسط يمكن إعادة كتابة الاحتمال الشرطي:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

A, B = الحدث

P(A|B) = احتمال A يعطي B

P(B|A) = احتمال B يعطي A

P(A), P(B) = احتمال كل من A و B

- الخطوات اللازمة لبناء خوارزمية بييز المبسط (Naïve Bayes)⁹⁵:

- 1- حساب احتمال كل فئة وفقاً لأمثلة التدريب.
- 2- حساب الاحتمال الشرطي لجميع الأقسام لكل سمة مميزة.
- 3- حساب ناتج احتمالية الفئة والاحتمال الشرطي لكل فئة.
- 4- تحديد العنصر ذي القيمة الأكبر في الخطوات الثلاث باعتباره فئة التصنيف.

- مميزات خوارزمية بييز المبسط (Naïve Bayes)⁹⁶:

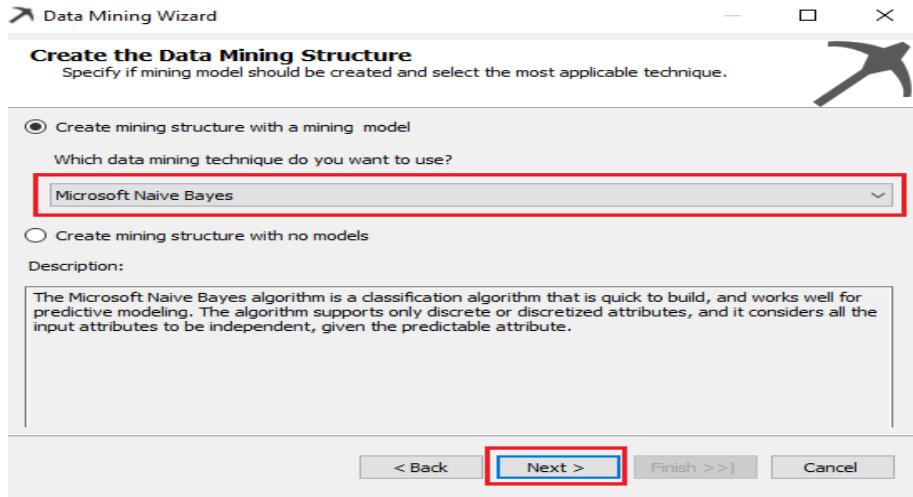
1. منطق الخوارزمية بسيط وسهل وسريع التنفيذ.
2. غير حساس للبيانات المفقودة.
3. حينما تكون الخصائص غير مرتبطة (مستقلة) يكون أداء الخوارزمية أفضل من الخوارزميات الأخرى.
4. يعمل بشكل أفضل مع البيانات الفئوية مقارنة مع البيانات الرقمية.

⁹⁵ /https://arabicprogrammer.com/article/94251567484

⁹⁶ /https://arabicprogrammer.com/article/94251567484

تطبيق بيز المبسط (Naïve Bayes) على برنامج Visual Studio 2019:

سيتم تطبيق بيز المبسط (Naïve Bayes) لأول مرة على بيانات عام 2018 والذي يحتوي على 1,224 سجل لعملاء حاصلين على تسهيلات، ومن ثم سيتم مقارنة النموذج مع العملاء الحاصلين على تسهيلات جديدة عام 2019 لنبين فيه دقة النموذج عند مقارنته مع بيانات واقعية، ومن ثم سيتم بناء نموذج ثاني يحتوي على بيانات عام 2018 و 2019 ومقارنته مع عام 2020، ومن ثم سيتم بناء نموذج نهائي يحتوي على بيانات 2018 و2019 و2020 .



شكل (30) اختيار خوارزمية بيز المبسط (Naïve Bayes) عن بيانات 2018

لبناء نموذج التنقيب نختار من 'Mining structure' Microsoft Naive Bayes كما هو موضح في الشكل (30)، ومن ثم نقوم بتحديد الواصفات المراد تحديدها كإدخال (input) للخوارزمية والواصفات المراد استبعادها والواصفات المراد التنبؤ بها كما تم شرحها في خوارزمية شجر القرار.

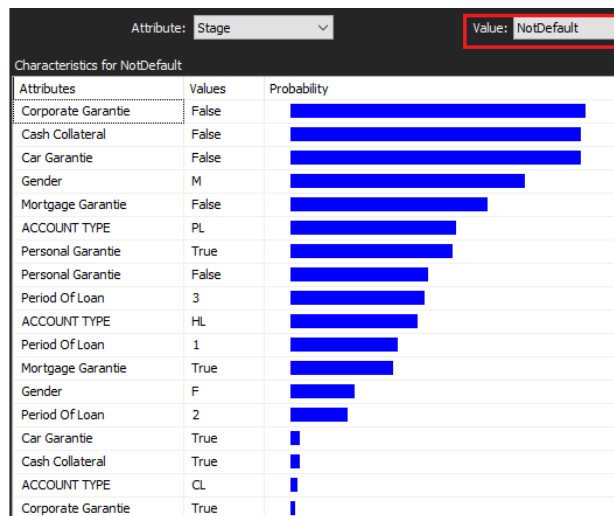
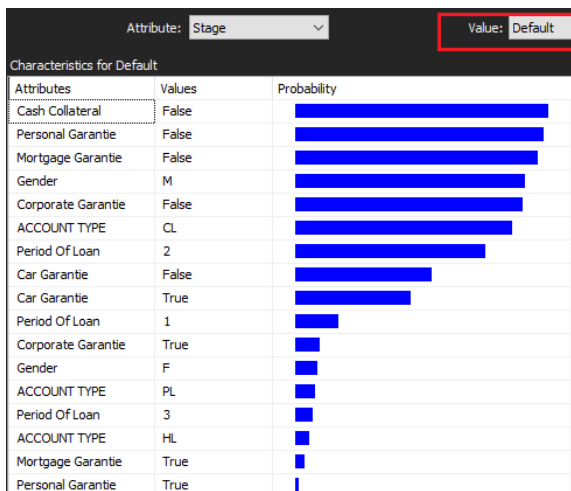
| Attributes | States | Population (All) Size: 857 | NotDefault Size: 253 | Default Size: 604 | missing Size: 0 |
|--------------------|---------------------------|-------------------------------|-------------------------|----------------------|--------------------|
| ACCOUNT TYPE | CL PL HL Missing | | | | |
| Car Garantie | False True Missing | | | | |
| Cash Collateral | False True Missing | | | | |
| Corporate Garantie | False True Missing | | | | |
| Gender | M F Missing | | | | |
| Mortgage Garantie | False True Missing | | | | |
| Period Of Loan | 2 1 3 Missing | | | | |
| Personal Garantie | False True Missing | | | | |

شكل (31) ملف السمات عن بيانات 2018

| Attributes | States | Population (All) | NotDefault | Default | missing |
|--------------------|---------|------------------|------------|---------|---------|
| Size | | 857 | 253 | 604 | 0 |
| ACCOUNT TYPE | CL | 525 | 2% | 86% | 0 |
| ACCOUNT TYPE | PL | 190 | 55% | 8% | 0 |
| ACCOUNT TYPE | HL | 142 | 42% | 6% | 0 |
| ACCOUNT TYPE | Missing | 0 | 0% | 0% | 0 |
| Car Garantie | FALSE | 573 | 97% | 54% | 0 |
| Car Garantie | TRUE | 284 | 3% | 46% | 0 |
| Car Garantie | Missing | 0 | 0% | 0% | 0 |
| Cash Collateral | FALSE | 849 | 97% | 100% | 0 |
| Cash Collateral | TRUE | 8 | 3% | 0% | 0 |
| Cash Collateral | Missing | 0 | 0% | 0% | 0 |
| Corporate Garantie | FALSE | 792 | 98% | 90% | 0 |
| Corporate Garantie | TRUE | 65 | 2% | 10% | 0 |
| Corporate Garantie | Missing | 0 | 0% | 0% | 0 |
| Gender | M | 749 | 78% | 91% | 0 |
| Gender | F | 107 | 21% | 9% | 0 |
| Gender | Missing | 0 | 0% | 0% | 0 |
| Gender | 1 | 1 | 0% | 0% | 0 |
| Mortgage Garantie | FALSE | 745 | 66% | 96% | 0 |
| Mortgage Garantie | TRUE | 112 | 34% | 4% | 0 |
| Mortgage Garantie | Missing | 0 | 0% | 0% | 0 |
| Period Of Loan | 2 | 504 | 19% | 75% | 0 |
| Period Of Loan | 1 | 196 | 36% | 17% | 0 |
| Period Of Loan | 3 | 157 | 45% | 7% | 0 |
| Period Of Loan | Missing | 0 | 0% | 0% | 0 |
| Personal Garantie | FALSE | 709 | 46% | 98% | 0 |
| Personal Garantie | TRUE | 148 | 54% | 2% | 0 |
| Personal Garantie | Missing | 0 | 0% | 0% | 0 |

جدول (11) جدول تفصيلي لملف السمات بالنسب عن بيانات 2018

من خلال تطبيق خوارزمية بيبز المبسط يمكن الحصول على احتمالية التعثر أو عدم التعثر بناء على كل سمة كما هو موضح بالشكل (31) وكما تم شرحها بالجدول (11)، وقد تم تدريب 70% من بيانات 2018 والتي تتمثل بـ 1,224 سجل واختبار 30% منها.



شكل (32) احتماليات التعثر وعدم التعثر وفق خصائص السمات لبيانات 2018

من خلال تطبيق خوارزمية بيز المبسط يمكن الحصول أيضاً على شكل بياني يظهر فيه احتمالية التعثر وعدم التعثر وفق خصائص السمات التي تم أخذها بالاعتبار كما هو موضح بالشكل (32).

| Predicted | NotDefault (Actual) | Default (Actual) |
|------------|---------------------|------------------|
| NotDefault | 104 | 31 |
| Default | 10 | 222 |

جدول (12) مصفوفة التعارض لخوارزمية بيز المبسط عن بيانات 2018

يبين الجدول (12) مصفوفة التعارض وقد تم اختبار 30% من البيانات المدخلة عن عام 2018 والتي تمثل 367 سجل من أصل 1224 سجل.

وكما تبين عدد السجلات التي كانت في الواقع غير متعثرين وتم التوقع بعدم تعثرهم 104 سجل ما يعادل 28.3% من العينة، وعدد السجلات في الواقع هم غير متعثرين وتم التنبؤ بتعثرهم 10 سجل أي ما يعادل 2.7%، وعدد السجلات التي في الواقع متعثرين وتم التنبؤ بعدم تعثرهم 31 سجل ما يعادل 8.4%، بينما تم التوقع بتعثر 222 سجل وهم في الواقع متعثرين أيضاً أي ما يعادل 60.5%، ومنه نستنتج بأن نسبة دقة النموذج بعد اختبار جزء من العينة المدخلة 88.8%.

| ACCOUNT | ACCOUNT | AgnetAge | CarGarant | CashColla | Corporate | Gender | ID | IsOtherBa | Mortgage | period_of | PeriodOfL | PersonalG | PROFESSI | stage | Predicatic | Probabilit | RESULT |
|---------|---------|----------|-----------|-----------|-----------|--------|-----|-----------|----------|-----------|-----------|-----------|----------|------------|------------|------------|--------|
| RETAIL | PL | 2 | FALSE | FALSE | FALSE | F | 143 | FALSE | FALSE | 1-M | 1 | TRUE | employee | NotDefault | NotDefau | 0.996459 | TRUE |
| RETAIL | PL | 2 | FALSE | FALSE | FALSE | M | 144 | FALSE | FALSE | 1-M | 1 | TRUE | employee | NotDefault | NotDefau | 0.994235 | TRUE |
| RETAIL | PL | 2 | FALSE | FALSE | FALSE | M | 145 | FALSE | FALSE | 1-M | 1 | TRUE | employee | NotDefault | NotDefau | 0.994235 | TRUE |
| RETAIL | PL | 2 | FALSE | FALSE | FALSE | M | 146 | FALSE | FALSE | 1-M | 1 | TRUE | employee | NotDefault | NotDefau | 0.994235 | TRUE |
| RETAIL | PL | 2 | FALSE | FALSE | FALSE | M | 147 | FALSE | FALSE | 1-M | 1 | TRUE | employee | NotDefault | NotDefau | 0.994235 | TRUE |
| RETAIL | PL | 2 | FALSE | FALSE | FALSE | F | 148 | FALSE | FALSE | 1-M | 1 | TRUE | employee | NotDefault | NotDefau | 0.996459 | TRUE |
| RETAIL | PL | 2 | FALSE | FALSE | FALSE | M | 149 | FALSE | FALSE | 1-M | 1 | TRUE | employee | NotDefault | NotDefau | 0.994235 | TRUE |
| RETAIL | PL | 2 | FALSE | FALSE | FALSE | M | 150 | FALSE | FALSE | 1-M | 1 | TRUE | employee | NotDefault | NotDefau | 0.994235 | TRUE |
| RETAIL | PL | 2 | FALSE | FALSE | FALSE | M | 177 | FALSE | FALSE | 1-M | 1 | TRUE | employee | Default | NotDefau | 0.994235 | FALSE |
| RETAIL | PL | 2 | FALSE | FALSE | FALSE | M | 152 | FALSE | FALSE | 1-M | 1 | TRUE | employee | NotDefault | NotDefau | 0.994235 | TRUE |
| RETAIL | PL | 2 | FALSE | FALSE | FALSE | M | 153 | FALSE | FALSE | 1-M | 1 | TRUE | employee | NotDefault | NotDefau | 0.994235 | TRUE |
| RETAIL | PL | 2 | FALSE | FALSE | FALSE | M | 154 | FALSE | FALSE | 1-M | 1 | TRUE | employee | NotDefault | NotDefau | 0.994235 | TRUE |
| RETAIL | PL | 2 | FALSE | FALSE | FALSE | F | 155 | FALSE | FALSE | 1-M | 1 | TRUE | employee | NotDefault | NotDefau | 0.996459 | TRUE |
| RETAIL | PL | 2 | FALSE | FALSE | FALSE | F | 156 | FALSE | FALSE | 1-M | 1 | TRUE | employee | NotDefault | NotDefau | 0.996459 | TRUE |
| RETAIL | PL | 2 | FALSE | FALSE | FALSE | M | 157 | FALSE | FALSE | 1-M | 1 | TRUE | employee | NotDefault | NotDefau | 0.994235 | TRUE |
| RETAIL | PL | 2 | FALSE | FALSE | FALSE | F | 158 | FALSE | FALSE | 1-M | 1 | TRUE | employee | NotDefault | NotDefau | 0.996459 | TRUE |
| RETAIL | PL | 2 | FALSE | FALSE | FALSE | M | 159 | FALSE | FALSE | 1-M | 1 | TRUE | employee | NotDefault | NotDefau | 0.994235 | TRUE |
| RETAIL | PL | 2 | FALSE | FALSE | FALSE | M | 160 | FALSE | FALSE | 1-M | 1 | TRUE | employee | NotDefault | NotDefau | 0.994235 | TRUE |
| RETAIL | PL | 2 | FALSE | FALSE | FALSE | F | 161 | FALSE | FALSE | 1-M | 1 | TRUE | employee | NotDefault | NotDefau | 0.996459 | TRUE |
| RETAIL | PL | 2 | FALSE | FALSE | FALSE | F | 162 | FALSE | FALSE | 1-M | 1 | TRUE | employee | NotDefault | NotDefau | 0.996459 | TRUE |
| RETAIL | PL | 2 | FALSE | FALSE | FALSE | M | 163 | FALSE | FALSE | 1-M | 1 | TRUE | employee | NotDefault | NotDefau | 0.994235 | TRUE |

جدول (13) مقارنة العملاء الجدد 2019 مع نتائج خوارزمية بيز المبسط

تم تطبيق خوارزمية بيز المبسط على العملاء الحاصلين على تسهيلات جديدة خلال عام 2019 وتم الحصول على نسبة احتمال حدوث التعثر أو عدم التعثر كما هو وارد في الجدول (12) وكانت النتائج وفق ما هو وارد في الجدول (14):

| predicted | Default(Actual) | NotDefault(Actual) |
|------------|-----------------|--------------------|
| Default | 0 | 5 |
| NotDefault | 3 | 516 |

جدول (14) نتائج مقارنة العملاء الجدد 2019 مع ببيز المبسط

ومنه نستنتج بأن درجة دقة النموذج مقارنة مع العملاء الجدد 2019 هي 98.5% أي تم التنبؤ ب 516 سجل من أصل 524 بينما لم يتمكن النموذج بالتنبؤ بشكل صحيح بنسبة 1.5%.

بعد الانتهاء من بناء النموذج الأول لبييز المبسط بناء على بيانات 2018 واستنباط القواعد وتطبيقها على بيانات 2019 واستنتاج درجة دقة النموذج، سيتم تطبيق النموذج الثاني لبييز المبسط بناء على بيانات 2018 و2019 بشكل كامل واستخراج القواعد وتطبيقها على بيانات 2020 لاستنتاج درجة دقة النموذج، وسيتم تطبيق نفس خطوات بناء النموذج الأول.

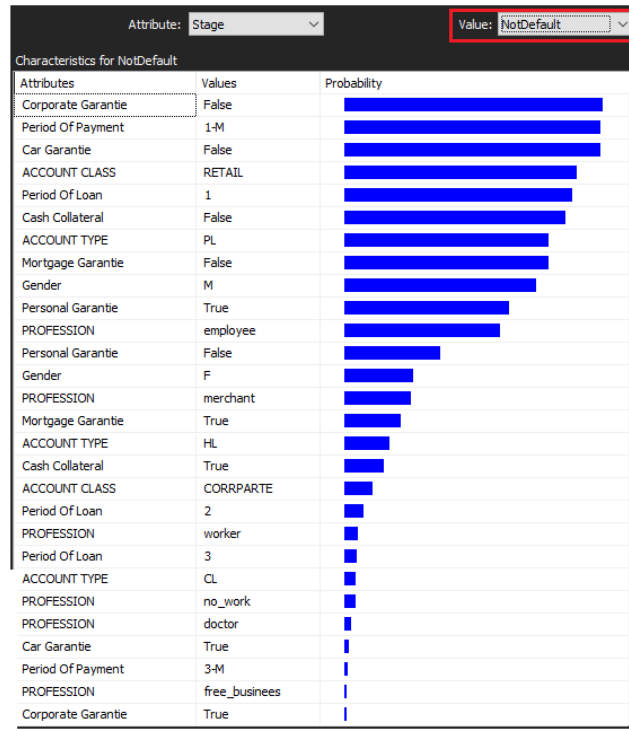
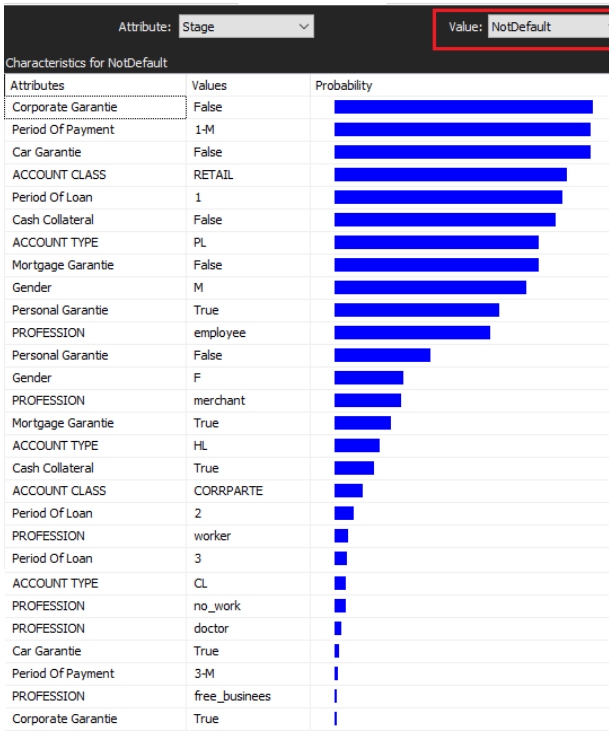


شكل (33) ملف السمات عن بيانات 2018 و2019

| Attributes | States | Population (All) | NotDefault | Default | missing |
|--------------------|----------------|------------------|------------|---------|---------|
| Size | | 1246 | 646 | 600 | 0 |
| ACCOUNT CLASS | RETAIL | 1147 | 89% | 96% | 0 |
| ACCOUNT CLASS | CORRPORTE | 99 | 11% | 5% | 0 |
| ACCOUNT CLASS | Missing | 0 | 0% | 0% | 0 |
| ACCOUNT TYPE | PL | 574 | 78% | 11% | 0 |
| ACCOUNT TYPE | CL | 519 | 4% | 82% | 0 |
| ACCOUNT TYPE | HL | 153 | 17% | 7% | 0 |
| ACCOUNT TYPE | Missing | 0 | 0% | 0% | 0 |
| Car Garantie | FALSE | 943 | 98% | 52% | 0 |
| Car Garantie | TRUE | 303 | 2% | 49% | 0 |
| Car Garantie | Missing | 0 | 0% | 0% | 0 |
| Cash Collateral | FALSE | 1147 | 85% | 100% | 0 |
| Cash Collateral | TRUE | 99 | 15% | 0% | 0 |
| Cash Collateral | Missing | 0 | 0% | 0% | 0 |
| Corporate Garantie | FALSE | 1176 | 99% | 89% | 0 |
| Corporate Garantie | TRUE | 70 | 1% | 11% | 0 |
| Corporate Garantie | Missing | 0 | 0% | 0% | 0 |
| Gender | M | 1025 | 74% | 92% | 0 |
| Gender | F | 221 | 27% | 8% | 0 |
| Gender | Missing | 0 | 0% | 0% | 0 |
| Mortgage Garantie | FALSE | 1091 | 78% | 98% | 0 |
| Mortgage Garantie | TRUE | 155 | 22% | 3% | 0 |
| Mortgage Garantie | Missing | 0 | 0% | 0% | 0 |
| Period Of Loan | 1 | 838 | 88% | 46% | 0 |
| Period Of Loan | 2 | 344 | 7% | 49% | 0 |
| Period Of Loan | 3 | 64 | 5% | 5% | 0 |
| Period Of Loan | Missing | 0 | 0% | 0% | 0 |
| Period Of Payment | 1-M | 1233 | 98% | 100% | 0 |
| Period Of Payment | 3-M | 12 | 2% | 0% | 0 |
| Period Of Payment | 2-M | 0 | 0% | 0% | 0 |
| Period Of Payment | 6-M | 1 | 0% | 0% | 0 |
| Period Of Payment | Missing | 0 | 0% | 0% | 0 |
| Personal Garantie | FALSE | 826 | 37% | 98% | 0 |
| Personal Garantie | TRUE | 420 | 63% | 2% | 0 |
| Personal Garantie | Missing | 0 | 0% | 0% | 0 |
| PROFESSION | employee | 590 | 60% | 34% | 0 |
| PROFESSION | merchant | 435 | 26% | 45% | 0 |
| PROFESSION | worker | 100 | 6% | 11% | 0 |
| PROFESSION | free_busineses | 39 | 1% | 5% | 0 |
| PROFESSION | no_work | 36 | 4% | 1% | 0 |
| PROFESSION | doctor | 34 | 3% | 3% | 0 |
| PROFESSION | lawyer | 6 | 0% | 1% | 0 |
| PROFESSION | retired | 4 | 0% | 0% | 0 |

جدول (15) جدول تفصيلي لملف السمات بالنسب عن بيانات 2018 و2019

وقد تم تدريب 70% من بيانات 2018 و2019 التي تتمثل بـ 1,779 سجل واختبار 30% منها، وظهرت ملف السمات التي يعتمد عليها النموذج لبناء احتمالات التعثر كما هو وارد في الشكل (33) والجدول (15).



شكل (34) احتماليات التعثر وعدم التعثر وفق خصائص السمات لبيانات 2018 و2019

من خلال تطبيق خوارزمية بيز الميسر يمكن الحصول أيضاً على شكل بياني يظهر فيه احتمالية التعثر وعدم التعثر وفق خصائص السمات التي تم أخذها بالاعتبار كما هو موضح بالشكل(33).

| Predicted | NotDefault (Actual) | Default (Actual) |
|------------|---------------------|------------------|
| NotDefault | 250 | 23 |
| Default | 19 | 241 |

جدول (16) مصفوفة التعارض لخوارزمية بيز الميسر عن بيانات 2018 و2019

يبين الجدول (16) مصفوفة التعارض وقد تم اختبار 30% من البيانات المدخلة عن عامي 2018 و2019 والتي تمثل 533 سجل من أصل 1779 سجل.

وكما تبين عدد السجلات التي كانت في الواقع غير متعثرين وتم التوقع بعدم تعثرهم 250 سجل ما يعادل 46.9% من العينة، وعدد السجلات في الواقع هم غير متعثرين وتم التنبؤ بتعثرهم 19 سجل أي ما يعادل 3.6%، وعدد السجلات التي في الواقع متعثرين وتم التنبؤ بعدم تعثرهم 23 سجل ما يعادل 4.3%، بينما تم التوقع بتعثر 241 سجل وهم في الواقع متعثرين أيضاً أي ما يعادل 45.2%، ومنه نستنتج بأن نسبة دقة النموذج بعد اختبار جزء من العينة المدخلة 92.1%.

| ACCOUNT | ACCOUNT | AgnetAge | CarGarant | CashColla | Corporate | Gender | ID | IsOtherBa | LIMITAMC | Mortgage | period_of | PeriodOfL | PersonalG | PROFESSI | stage | Predicatic | Probabilit | Result |
|---------|---------|----------|-----------|-----------|-----------|--------|----|-----------|----------|----------|-----------|-----------|-----------|----------|----------|------------|------------|--------|
| RETAIL | PL | 2 | TRUE | FALSE | FALSE | M | 1 | TRUE | 0 | FALSE | 1-M | 3 | FALSE | employee | NotDefau | Default | 0.872023 | FALSE |
| RETAIL | HL | 3 | FALSE | FALSE | FALSE | M | 2 | FALSE | 0 | TRUE | 1-M | 3 | FALSE | merchant | NotDefau | NotDefau | 0.862908 | TRUE |
| RETAIL | PL | 2 | FALSE | FALSE | FALSE | M | 3 | FALSE | 0 | FALSE | 1-M | 3 | TRUE | employee | NotDefau | NotDefau | 0.99597 | TRUE |
| RETAIL | PL | 2 | TRUE | FALSE | FALSE | M | 4 | TRUE | 0 | FALSE | 1-M | 3 | TRUE | employee | NotDefau | NotDefau | 0.922735 | TRUE |
| RETAIL | PL | 2 | TRUE | FALSE | FALSE | M | 5 | FALSE | 0 | FALSE | 1-M | 3 | TRUE | employee | NotDefau | NotDefau | 0.890367 | TRUE |
| RETAIL | HL | 2 | FALSE | FALSE | FALSE | M | 6 | FALSE | 0 | TRUE | 1-M | 3 | FALSE | employee | NotDefau | NotDefau | 0.949306 | TRUE |
| RETAIL | HL | 2 | FALSE | FALSE | FALSE | M | 7 | FALSE | 0 | FALSE | 1-M | 3 | FALSE | employee | NotDefau | NotDefau | 0.640906 | TRUE |
| RETAIL | PL | 2 | FALSE | FALSE | FALSE | M | 8 | FALSE | 0 | FALSE | 1-M | 3 | TRUE | employee | NotDefau | NotDefau | 0.99597 | TRUE |
| RETAIL | PL | 2 | FALSE | FALSE | FALSE | M | 9 | FALSE | 0 | FALSE | 1-M | 3 | TRUE | employee | NotDefau | NotDefau | 0.99597 | TRUE |
| RETAIL | PL | 2 | FALSE | FALSE | FALSE | F | 10 | FALSE | 0 | FALSE | 1-M | 3 | TRUE | employee | NotDefau | NotDefau | 0.997784 | TRUE |
| RETAIL | PL | 2 | FALSE | FALSE | FALSE | M | 11 | FALSE | 0 | TRUE | 1-M | 3 | FALSE | employee | NotDefau | NotDefau | 0.979797 | TRUE |
| RETAIL | PL | 2 | FALSE | FALSE | FALSE | F | 12 | FALSE | 0 | FALSE | 1-M | 3 | TRUE | employee | NotDefau | NotDefau | 0.997784 | TRUE |
| RETAIL | PL | 2 | FALSE | FALSE | FALSE | M | 13 | FALSE | 0 | TRUE | 1-M | 3 | TRUE | employee | NotDefau | NotDefau | 0.998173 | TRUE |
| RETAIL | PL | 2 | FALSE | FALSE | FALSE | F | 14 | FALSE | 0 | FALSE | 1-M | 3 | TRUE | employee | NotDefau | NotDefau | 0.997784 | TRUE |
| RETAIL | PL | 2 | FALSE | FALSE | FALSE | M | 15 | FALSE | 0 | TRUE | 1-M | 3 | TRUE | employee | NotDefau | NotDefau | 0.998173 | TRUE |
| RETAIL | PL | 2 | FALSE | FALSE | FALSE | M | 16 | FALSE | 0 | FALSE | 1-M | 3 | TRUE | employee | NotDefau | NotDefau | 0.99597 | TRUE |
| RETAIL | PL | 2 | FALSE | FALSE | FALSE | M | 17 | FALSE | 0 | FALSE | 1-M | 3 | TRUE | employee | NotDefau | NotDefau | 0.99597 | TRUE |
| RETAIL | PL | 2 | TRUE | FALSE | FALSE | M | 18 | FALSE | 0 | FALSE | 1-M | 3 | FALSE | employee | NotDefau | Default | 0.909337 | FALSE |
| RETAIL | PL | 2 | FALSE | FALSE | FALSE | M | 19 | FALSE | 0 | FALSE | 1-M | 3 | TRUE | employee | NotDefau | NotDefau | 0.99597 | TRUE |
| RETAIL | HL | 2 | FALSE | FALSE | FALSE | M | 20 | FALSE | 0 | FALSE | 1-M | 3 | FALSE | employee | NotDefau | NotDefau | 0.640906 | TRUE |
| RETAIL | PL | 2 | FALSE | FALSE | FALSE | M | 21 | FALSE | 0 | FALSE | 1-M | 3 | TRUE | employee | NotDefau | NotDefau | 0.99597 | TRUE |

جدول (17) مقارنة العملاء الجدد 2020 مع نتائج خوارزمية بيز المبسط

تم تطبيق خوارزمية بيز المبسط على العملاء الحاصلين على تسهيلات جديدة خلال عام 2020 وتم الحصول على نسبة احتمال حدوث التعثر أو عدم التعثر كما هو وارد في الجدول (17) وكانت النتائج وفق ما هو وارد في الجدول (18):

| predicted | Default(Actual) | NotDefault(Actual) |
|------------|-----------------|--------------------|
| Default | 4 | 5 |
| NotDefault | 12 | 187 |

جدول (18) نتائج مقارنة العملاء الجدد 2020 مع بيز المبسط

ومنه نستنتج بأن درجة دقة النموذج مقارنة مع العملاء الجدد 2020 هي 91.8% أي تم التنبؤ ب 191 سجل من أصل 208 بينما لم يتمكن النموذج بالتنبؤ بشكل صحيح بنسبة 8.2%.

بعد الانتهاء من بناء النموذج الثاني لبيز المبسط بناء على بيانات 2018 و 2019 واستنباط القواعد وتطبيقها على بيانات 2020 واستنتاج درجة دقة النموذج، سيتم تطبيق النموذج النهائي من بيز المبسط بناء على بيانات 2018 و 2019 و 2020 بشكل كامل واستخراج القواعد، وسيتم تطبيق نفس خطوات بناء النموذج الأول والثاني.

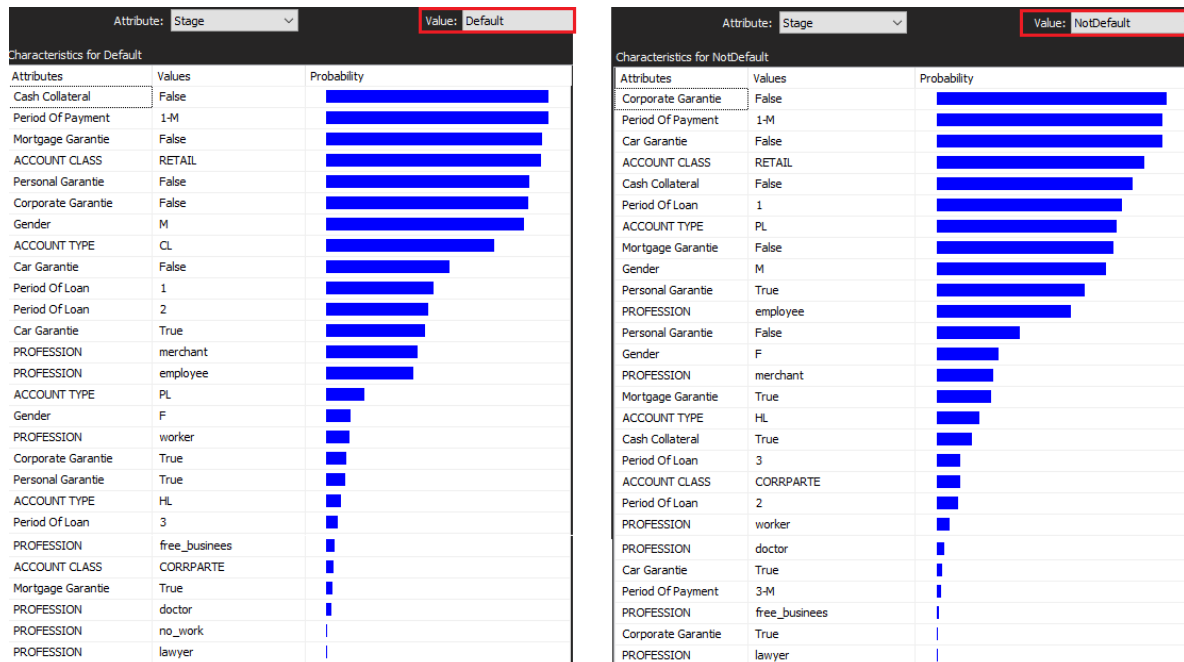


شكل (35) ملف السمات عن بيانات 2018 و2019 و2020

| Attributes | States | Population (All) | NotDefault | Default | missing |
|--------------------|-----------------|------------------|------------|---------|---------|
| Size | | 1789 | 948 | 841 | 0 |
| ACCOUNT CLASS | RETAIL | 1659 | 90% | 96% | 0 |
| ACCOUNT CLASS | CORRPARTE | 130 | 10% | 4% | 0 |
| ACCOUNT CLASS | Missing | 0 | 0% | 0% | 0 |
| ACCOUNT TYPE | PL | 880 | 78% | 17% | 0 |
| ACCOUNT TYPE | CL | 671 | 4% | 76% | 0 |
| ACCOUNT TYPE | HL | 238 | 19% | 7% | 0 |
| ACCOUNT TYPE | Missing | 0 | 0% | 0% | 0 |
| Car Garantie | FALSE | 1391 | 98% | 56% | 0 |
| Car Garantie | TRUE | 398 | 3% | 45% | 0 |
| Car Garantie | Missing | 0 | 0% | 0% | 0 |
| Cash Collateral | FALSE | 1642 | 85% | 100% | 0 |
| Cash Collateral | TRUE | 147 | 15% | 0% | 0 |
| Cash Collateral | Missing | 0 | 0% | 0% | 0 |
| Corporate Garantie | FALSE | 1703 | 99% | 91% | 0 |
| Corporate Garantie | TRUE | 86 | 1% | 9% | 0 |
| Corporate Garantie | Missing | 0 | 0% | 0% | 0 |
| Gender | M | 1442 | 73% | 89% | 0 |
| Gender | F | 347 | 27% | 11% | 0 |
| Gender | Missing | 0 | 0% | 0% | 0 |
| Mortgage Garantie | FALSE | 1536 | 76% | 97% | 0 |
| Mortgage Garantie | TRUE | 253 | 24% | 3% | 0 |
| Mortgage Garantie | Missing | 0 | 0% | 0% | 0 |
| Period Of Loan | 1 | 1168 | 80% | 49% | 0 |
| Period Of Loan | 2 | 475 | 9% | 46% | 0 |
| Period Of Loan | 3 | 146 | 11% | 6% | 0 |
| Period Of Loan | Missing | 0 | 0% | 0% | 0 |
| Period Of Payment | 1-M | 1765 | 98% | 100% | 0 |
| Period Of Payment | 3-M | 23 | 2% | 0% | 0 |
| Period Of Payment | 6-M | 1 | 0% | 0% | 0 |
| Period Of Payment | 2-M | 0 | 0% | 0% | 0 |
| Period Of Payment | Missing | 0 | 0% | 0% | 0 |
| Personal Garantie | FALSE | 1108 | 36% | 91% | 0 |
| Personal Garantie | TRUE | 681 | 64% | 9% | 0 |
| Personal Garantie | Missing | 0 | 0% | 0% | 0 |
| PROFESSION | employee | 884 | 58% | 40% | 0 |
| PROFESSION | merchant | 579 | 25% | 41% | 0 |
| PROFESSION | worker | 146 | 6% | 11% | 0 |
| PROFESSION | no_work | 61 | 6% | 1% | 0 |
| PROFESSION | doctor | 55 | 4% | 3% | 0 |
| PROFESSION | free_businesses | 46 | 1% | 4% | 0 |
| PROFESSION | lawyer | 11 | 1% | 1% | 0 |
| PROFESSION | consultant | 4 | 0% | 0% | 0 |

جدول (19) جدول تفصيلي لملف السمات بالنسبة عن بيانات 2018 و2019 و2020

وقد تم تدريب 90% من بيانات 2018 و2019 و2020 التي تتمثل بـ 1987 سجل واختبار 10% من البيانات بدلاً من 30% كما هو وارد بالنموذج الأول والثاني وذلك لأننا بحاجة الى تدريب أكبر قدر ممكن من البيانات والتحقق من درجة دقة النموذج من خلال 10% من البيانات المتبقية، وظهرت ملف السمات التي يعتمد عليها النموذج لبناء احتمالات التعثر كما هو وارد في الشكل (35) والجدول (18).



شكل (36) احتماليات التعثر وعدم التعثر وفق خصائص السمات لبيانات 2018 و2019 و2020

من خلال تطبيق خوارزمية ببيز المبسط يمكن الحصول أيضاً على شكل بياني يظهر فيه احتمالية التعثر وعدم التعثر وفق خصائص السمات التي تم أخذها بالاعتبار كما هو موضح بالشكل (36).

| Predicted | NotDefault (Actual) | Default (Actual) |
|------------|---------------------|------------------|
| NotDefault | 90 | 16 |
| Default | 8 | 84 |

جدول (20) مصفوفة التعارض لخوارزمية ببيز المبسط عن بيانات 2018 و2019 و2020

يبين الجدول (20) مصفوفة التعارض وقد تم اختبار 10% من البيانات المدخلة عن الأعوام 2018 و2019 و2020 والتي تمثل 198 سجل من أصل 1987 سجل.

وكما تبين عدد السجلات التي كانت في الواقع غير متعثرين وتم التوقع بعدم تعثرهم 90 سجل ما يعادل 45.5% من العينة، وعدد السجلات في الواقع هم غير متعثرين وتم التنبؤ بتعثرهم 8 سجل أي ما يعادل 4%، وعدد السجلات التي في الواقع متعثرين وتم التنبؤ بعدم تعثرهم 16 سجل ما يعادل 8%، بينما تم التوقع بتعثر 84 سجل وهم في الواقع متعثرين أيضاً أي ما يعادل 42.5%، ومنه نستنتج بأن نسبة دقة النموذج بعد اختبار جزء من العينة المدخلة 87.9%.

الفصل الرابع: النتائج والتوصيات

أولاً- النتائج:

- 1- من خلال تطبيق خوارزمية شجرة القرار تم الحصول على شجرة سهلة الفهم والقراءة وتطبيقها يساعد متخذي القرار على توقعات مستقبلية فيستطيع من خلالها اتخاذ إجراءات احتياطية لمنع حدوثها وكان عامل "نوع التمويل" هو الأول (جذر الشجرة) في عملية التوقع.
- 2- من خلال القواعد التي تم استنتاجها من شجرة القرار تساعد متخذي قرار المنح الائتماني إما برفض أو قبول طلبات المنح الائتماني أو بتعديل بشروط المنح، كما تساعد على تشديد المتابعة للعملاء المنتبئ بتعثرهم.
- 3- تُمكن خوارزمية بيبز المبسط في الحصول على الاحتماليات بالنسب والذي يساعد متخذ القرار في معرفة اقتراب النسبة من حالة التعثر أم عدم التعثر.
- 4- تقارب درجة دقة النموذج وكما أعطت نسب عالية من الدقة عند مقارنتها مع بيانات واقعية.
- 5- أفضل وأدق النتائج ظهرت باستخدام خوارزمية شجرة القرار.
- 6- النموذج المقترح وفر إطار عمل ساعد في التنبؤ باكتشاف المخاطر التي تهدد البنك وذلك من خلال تحديد الخصائص التي تشير الى وجود مخاطر تحد من منح العميل القرض.
- 7- استخدام تقنيات التنقيب في البيانات ساعد متخذ القرار في البنك من تحليل المخاطر المتعلقة من منح القروض والتنبؤ بها قبل وقوعها.
- 8- يوفر النموذج الكثير من الوقت والجهد على إدارة البنك في عملية دراسة القروض وتحليلها.
- 9- تحسن درجة دقة النموذج عند زيادة البيانات التاريخية المدخلة مما يدل الى الحاجة للحصول على بيانات تاريخية أكثر لتحسين التوقع ودرجة الدقة.

ثانياً- التوصيات:

- 1- استخدام خوارزميات أخرى في تنقيب البيانات وتطبيقها على مجموعة البيانات مثل K-means.
- 2- تطبيق خوارزميات تنقيب البيانات باستخدام برامج أخرى للتنقيب مثل برامج Orange و weka ومقارنة النتائج.
- 3- زيادة حجم البيانات المستخدمة.
- 4- إضافة بعض الخصائص المهمة التي يمكن أن تساعد في استخراج المزيد من المعرفة.

- 5- التركيز على تدريب وتطوير متخذي القرار على مثل هذه التقنيات مما يسهل عليهم في التوصل الى المعلومات المطلوبة في اتخاذ قراراتهم.
- 6- تطوير النموذج وربطه بمستودع للبيانات حتى يكون متاح لكل البنوك والمصارف.
- 7- بناء تطبيق ويب سهل الاستخدام بالنسبة للموظفين ومتخذي القرار.

قائمة المراجع:

أولاً: المراجع العربية:

- 1- أبو القاسم أحمد، هبة.(2015): "دراسة أثر تطبيق تقنية التصنيف على قطاع التأمين (دراسة حالة شركة شيكان)" رسالة ماجستير غير منشورة، كلية الدراسات العليا، جامعة النيلين.
- 2- بابكر، مصطفى. (٢٠٠٧)، التحليل باستخدام شجرة القرار، المعهد العربي للتخطيط، الكويت.
- 3- بالعجوز، حسين.(2005).إدارة المخاطر البنكية والتحكم فيها.مداخلة مقدمة الى الملتقى الوطني حول المنظومة المصرفية في الالفية الثالثة،جامعة جيجل الجزائر.ص9.
- 4- بسيوني، عبد الحميد.(2017)، الحاسب واكتشاف المعرفة تنقيب البيانات،دار الكتب العلمية للنشر والتوزيع، القاهرة.
- 5- بن عمر، خالد.(2011):"دراسة النماذج الحديثة لقياس مخاطر الائتمان لدى البنوك التجارية".رسالة دكتوراه غير منشورة، كلية العلوم الاقتصادية والتجارية وعلوم التسيير، جامعة بومرداس.
- 6- جمعة، أحمد حلمي (٢٠١٢)، استخدام الشبكات الاصطناعية في اكتشاف الأخطاء الجوهرية في البيانات المالية : دراسة تطبيقية، بحث مقدم إلى المؤتمر العلمي السنوي الحادي عشر بعنوان "ذكاء الأعمال واقتصاد المعرفة"، كلية الاقتصاد والعلوم الإدارية، جامعة الزيتونة الأردنية ٢٣ – ٢٦ ابريل، ص ص. ٤٧٩ – ٥٠٥
- 7- حسن، هالة حسن محمود.(2015): " تعدين بيانات التمويل الأصغر باستخدام تقنيات التصنيف والعنقدة ". رسالة ماجستير غير منشورة، كلية الدراسات العليا، جامعة النيلين.
- 8- حسن، وضاح عبدالله؛فقيري،سلام عثمان.(2018):" تطبيق تقنية تعلم المكنائ للتعنبؤ بمخاطر التمويل البنكي – دراسة حالة بنك العمال الوطني ". بحث منشور. كلية علوم الحاسوب وتقانة المعلومات، جامعة الزعيم الأزهرى.
- 9- الخطيب ، منال .(2004)"تكلفة الإئتمان المصرفي وقياس مخاطره بالتطبيق على أحد المصارف التجارية السورية". رسالة ماجستير ، كلية الاقتصاد، جامعة حلب.
- 10- الخليفة، شبيكة هاشم الطيب.(2017): " استخدام تنقيب البيانات في القطاع المصرفي-دراسة حالة في المصارف السودانية". رسالة ماجستير غير منشورة، كلية الدراسات العليا، جامعة النيلين.
- 11- الدغيم، عبد العزيز و الأمين، ابراهيم و أنجرو، إيمان.(2006)" التحليل الائتماني ودوره في ترشيد عمليات الإقراض المصرفي"، مجلة جامعة تشرين للدراسات والبحوث العلمية، سلسلة العلوم الاقتصادية والقانونية المجلة، العدد الثالث، ص 28.

- 12- الدوري، زكريا مطلق، أحمد، داليا عبد الحسين (2007)، " دور تنقيب البيانات في زيادة أداء المنظمة-دراسة تحليلية في المصرف الصناعي العراقي"، مجلة العلوم الاقتصادية والإدارية، العدد 48، العراق.
- 13- نيب، سوزان سمير وزملاؤه. (2011) إدارة الائتمان. الطبعة الأولى، عمان، الأردن: دار الفكر، ص 17
- 14- رضا، عبد المعطي. جودة، محفوظ أحمد. (1999) " إدارة الائتمان". عمان: دار وائل للنشر والتوزيع.
- 15- الزبيدي، محمود. (2001). إدارة الائتمان المصرفي والتحليل الائتماني. الطبعة 1، عمان: دار الوراق للنشر والتوزيع. ص 104.
- 16- زميت، محمد (2005)، "النظام المصرفي الجزائري في مواجهة تحديات العولمة المالية". رسالة ماجستير، كلية الاقتصاد، جامعة يوسف بن خدة.
- 17- زهرة، لعروسي قرين. (2017): "دور إدارة مخاطر الائتمان المصرفي في اتخاذ القرارات الائتمانية لدى البنوك التجارية-دراسة مجموعة من البنوك التجارية الجزائرية". رسالة دكتوراه منشورة، كلية العلوم الاقتصادية والتجارية علوم التسيير، جامعة محمد بوضياف بالمسيلة.
- 18- سليمان، الحيلي كباشي ابراهيم. (2017): " تنقيب بيانات المحلات التجارية بواسطة قاعدة الارتباط". رسالة ماجستير غير منشورة، كلية الدراسات العليا، جامعة النيلين.
- 19- السيسي، صلاح الدين. (2004). قضايا مصرفية معاصرة في الائتمان-الضمانات المصرفية- الاعتمادات المستندية. القاهرة: دار الفكر العربي. ص 30.
- 20- السيد، أحمد فايز أحمد . (2015) دراسة بعنوان ادوات التنقيب عن البيانات مفتوحة المصدر (دراسة تحليلية تقييمية)).
- 21- شاهين، علي عبد الله. (2010) " مدخل عملي لقياس مخاطر الائتمان المصرفي في البنوك التجارية – دراسة تحليلية تطبيقية". مكتبة الجامعة الإسلامية.
- 22- الشواربي، عبد الحميد محمد. إدارة المخاطر الائتمانية من وجهتي النظر المصرفية والقانونية، منشأة المعارف، الإسكندرية، مصر، 2002، ص 552-553.
- 23- الشخيلي، هديل. (2012) "العوامل الرئيسية المحددة لقرار الائتمان المصرفي في البنوك التجارية الأردنية". رسالة ماجستير، كلية الأعمال، جامعة الشرق الأوسط.
- 24- طه، عمر هاشم. (2013) "دور سياسات منح الائتمان المصرفي في تقليل المخاطر وزيادة الأرباح (دراسة ميدانية في مصرف الشمال للتنمية والاستثمار)، مجلة جامعة الكركوك للعلوم الإدارية والاقتصادية، العدد 2، المجلد 3، ص 59.

- 25- طيار، أحسن. شلابي، عمار (٢٠٠٥)، التنقيب في البيانات واتخاذ القرارات ، مجلة كلية العلوم الاقتصادية وعلوم التسيير ، جامعة ٢٠ أوت ١٩٥٥ سكيكدة ، الجزائر، العدد الرابع، ص ٨٤ – ٩٤
- 26- عبد المجيد، أمل محمد عز الدين.(2018): " استخدام تقنيات تنقيب البيانات في تحليل ومعرفة تفاصيل عمليات التمويل الأصغر، دراسة حالة بنك فيصل الإسلامي السوداني". رسالة ماجستير غير منشورة، كلية الدراسات العليا، جامعة النيلين.
- 27- عبدالله، خالد أحمد. (٢٠٠٤)، " قواعد البيانات المتقدمة – مستودعات البيانات"، كلية علوم الحاسوب وتقانة المعلومات ،جامعة العلوم والتقانة، السودان، يناير، ص ١,٢٠
- 28- عبيد الله،فايزة.(2016)، دعم أساليب القياس واتخاذ القرار باستخدام أساليب التنقيب عن البيانات كمدخل لإدارة التكاليف الاستراتيجية و دعم القدرة التنافسية للشركات الصناعية، كلية التجارة، جامعة دمنهور، دورية، ص 9.
- 29- عبيد، مصطفى. (2020) ، ورقة عمل بعنوان: التنقيب في قواعد البيانات واستكشاف المعلومات المخبأة فيها – وصف خوارزمياتها وتطبيقاتها في المجتمع.
- 30- العلاق ،بشير عباس . الإدارة الرقمية: المجالات و التطبيقات، مركز الإمارات للدراسات و البحوث الإستراتيجية، ابوظبي، 2005، ص 84.
- 31- علي، بسام، وزملاؤه ". (2011) استخدام تنقيب البيانات في التنبؤ بظاهرة تسرب طلبة الكلية بالتطبيق علي كلية الإدارة والاقتصاد"،المجلة العراقية للعلوم الاحصائية عدد(18)،ص209-236
- 32- العلي، عبد الستار. قنديلجي ،عامر إبراهيم. العمري، غسان .المدخل إلى إدارة المعرفة، دار المسيرة للنشر و التوزيع و الطباعة، الطبعة الأولى، عمان، 2006، ص. 157.
- 33- غنيم، أحمد. صناعة قرارات الائتمان والتمويل في إطار الاستراتيجية الشاملة للبنك، بدون ذكر دار النشر، بدون ذكر البلد، 2002، ص67.
- 34- الكاشف، محمود يوسف (2000)، "مدخل مقترح لتطوير دور المعلومات المحاسبية في إطار المفهوم المتكامل للجودة الشاملة"، مجلة الإدارة العامة، مجلد 94، العدد 3 ، أكتوبر.
- 35- كراجه ، عبد الحليم. ربابعة ، علي. السكران، ياسر. مطر، موسى.(2006). الإدارة والتحليل المالي. الطبعة 1، عمان: دار صنعاء للنشر والتوزيع.
- 36- كراسنة، ابراهيم.(2006). أطر أساسية و معاصرة في الرقابة على البنوك و إدارة المخاطر. أبو ظبي: معهد السياسات الاقتصادية.
- 37- المبارك، البدوي سعد البدوي.(2017): "استخدام تقنيات التنقيب في البيانات لاستكشاف أنماط مؤثرات التحصيل الأكاديمي لطلاب المرحلة الثانوية"رسالة ماجستير غير منشورة، كلية الدراسات العليا، جامعة النيلين.

- 38- نصرالله، هبة مبارك موسي.(2017): " التنبؤ بسداد قروض المصارف باستخدام تنقيب البيانات - دراسة الحالة بالتطبيق في بنك فيصل الإسلامي". رسالة ماجستير غير منشورة، كلية الدراسات العليا، جامعة النيلين.
- 39- تعميم رقم 4/م ن لعام 2019 الخاص بتطبيق بإلزام المصارف العاملة في سورية على تطبيق معيار التقارير المالية الدولي رقم 9.

ثانياً: المراجع الأجنبية:

- 1- Alemu, Teninet Belay (2019) "Predicting bank credit risk using data mining Technique: the case of bank of Abyssinia", A thesis, faculty of informatics, ST. Mary's university.
- 2- Amidu ، Mohammed and Hinson ، Robert (2006). Credit Risk ، Capital Structure and Lending Decisions of Banks in Ghana ، Banks and Bank Systems ، Vol. 1 ، No.1 ، P.93
- 3- B. Desai and Anita Desai, "The Role of Data mining in Banking Sector", IBA Bulletin, 2004.
- 4- Banasik J., Crook J., and Thomas L. (2003) "Sample selection bias in credit scoring models" Journal of the Operational Research Society, Vol 54, P. 822–832.
- 5- Barry de Ville (2006), Decision Trees for Business Intelligence and Data Mining: Using SAS Enterprise Miner, SAS Institute Inc., Cary, NC, USA, 2006, p1
- 6- Basel Committee. (2000) "Principles for the Management of Credit Risk"
- 7- Cebenoyan ، A. Sinan and Strahan ، Philip E. (2004) "Risk management ، capital structure and Lending at banks" ، Journal of Banking & Finance ، Vol. 28 P.19-43.
- 8- David E Vince. (2003) "Financial analysis and decision making". McGraw-Hill Companies. USA
- 9- Dilnutt, (2002), " KM in practice: three contemporary case studies", International journal of accounting information systems, vol.3, issue 2,

- 10- Fayyad, U. Piatetsky-Shapiro, G. & Smyth, P. (1996), —From data mining to knowledge Discovery in database, American Association for Artificial Intelligence Press, Cambridge
- 11- Feyssia, Ketema (2018) " The use of data mining to predict the loan repayment risk: the case of Oromia credit and saving share company ", A thesis, college of natural science, Addis Ababa University.
- 12- G. Dong and J. Li. Efficient mining of emerging patterns: Discovering trends and differences. In Proc. 1999 Int. Conf. Knowledge Discovery and Data Mining (KDD'99), pp. 43–52, San Diego, CA, Aug. 1999.
- 13- Gahlaut, Archana. Tushar. and Singh, Prince Kumar (2017) " Prediction analysis of risky credit using Data mining classification models", 8th International Conference on Computing, Communication and Networking Technologies (ICCCNT), 5-3 July 2017.
- 14- Hamid, Aboobyda Jafar. and Ahmed, tarig Mohammed. (2016) " Developing prediction model of loan risk in banks using Data mining", Machine Learning and Applications: An International Journal (MLAIJ), Vol.3, No.1, March 2016.
- 15- Han, J. and Kamber, M. (2006). Data Mining: Concepts and Techniques. Second Edition, Morgan Kaufmann Publishers, San Francisco
- 16- Hand d., Mannila H., Smyth R., Principles of Data Mining, MIT Press, London, 2001, p. 01.
- 17- Harry Zhang "The Optimality of Naive Bayes". FLAIRS2004 conference. (available online: PDF (<http://www.cs.unb.ca/profs/hzhang/publications/FLAIRS04ZhangH.pdf>))
- 18- Koksall, G., Batmaz, I., & Testik, M. C. (2011). A Review of Data Mining Applications for Quality Improvement in Manufacturing Industry. Expert Systems with Applications 38, pp. 13448-13467.
- 19- MacLennan, J., Tang, Z., & Crivat, B. (2009). Data Mining with Microsoft SQL Server 2008. Indianapolis,, Indiana, United States of America: Wiley Publishing, Inc.

- 20- Madaan,Mehul.Kumar,Aniket.Keshri,Chirag.Jain,Rachna.and Nagrath ,Preeti (2021) " Loan default prediction using decision trees and random forest: A comparative study",A thesis, Department of Electronics and Communication Engineering, GGSIP University.
- 21- MadanLalBhasin, —Data Mining: A Competitive Tool in the Banking and Retail Industries||, The Chartered Accountant October 2006
- 22- Michael, J., A. Berry and Gordan S. Linoff, (2000), Mastering Data Mining. John Wiley & Sons,In
- 23- Patel,Bhoomi.Patil,Harshal.Hembram,Jovita.and Jaswal,shree(2020) "Loan default forecasting using data mining",International conference for emerging technology(INCET), Jun5-7,2020.
- 24- Pyle, D., 1999 Data Preparation for Data Mining Morgan Kaufmann Publisher.
- 25- S. Kotsiantis , E. Koumanakos , D. Tzelepis , V. Tampakas (2006), Financial Application of Neural Networks: Two Case Studies in Greece, Artificial Neural Networks - ICANN, 16th International Conference, Athens, Greece, September 10-14, PP. 672 – 681
- 26- S. S. Kaptan, N S Chobey, "Indian Banking in Electronic Era", Sarup and Sons, Edition 2002
- 27- Sharma, A. & Panigrahi, P. K. (February 2012). A Review of Financial Accounting Fraud Detection Based on Data Mining Techniques. International Journal of Computer Applications 39 (1), pp. 37-47.
- 28- Silltow J., (2006).Data mining tools and techniques. United Kingdom
- 29- Turban, Efraim & Leidner, Dorothy (2008), Information Technology for Management, 6th edition, john wiley & sons, NewJersy, P183-186
- 30- Witten H. Ian and Frank, E. (2005). Data Mining: Practical Machine Learning Tools and Techniques, Second Edition. Morgan Kaufmann Publishers, San Francisco

