

تصميم شبكة عصبونية التفاضلية CNN لكشف عمليات الاحتيال في المعاملات المالية التي تستخدم بطاقة الائتمان

نورا كويس *

(تاريخ الإيداع 2023 /10/25 – تاريخ النشر 2023 /12/4)

□ ملخص □

تشهد الخدمات المالية تطوراً متسارعاً في العصر الحديث، إذ يتم تبني التقنيات المتقدمة لتسهيل وتحسين عمليات التعامل المالي. تعد البطاقات الائتمانية واحدة من الأدوات الرئيسية في هذا السياق، إذ توفر وسيلة سهلة ومريحة للدفع والشراء عبر الإنترنت وفي المتاجر التجارية. ومع ذلك، يرتبط ارتفاع استخدام البطاقات الائتمانية بزيادة في حوادث الاحتيال المالي. يتمثل الاحتيال في الحصول على معلومات بطاقة الائتمان أو استغلال بيانات العميل بطرق غير قانونية وغير مشروعة. لمكافحة هذه الظاهرة المتزايدة، تعتمد شركات البطاقات الائتمانية ومقدمي الخدمات المالية على تكنولوجيا الذكاء الاصطناعي لحماية العملاء وتحقيق أمن أعلى في العمليات المالية. يعد الذكاء الاصطناعي أداة قوية في اكتشاف ومنع الاحتيال، يهدف هذا البحث الى العمل على إيجاد تقنية قادرة على كشف الاحتيال على بطاقات الائتمان بأعلى دقة ممكنة مع نسبة أمن مرتفعة بالاعتماد على تدريب مجموعة من خوارزميات الذكاء الصناعي لتصنيف العمليات الاحتيالية

يقدم البحث نموذج تعلم عميق قادر على التمييز بي العمليات الاحتيالية والحقيقية المنفذة باستخدام بطاقات الائتمان وقد اعتمدت هذه المنهجية علي :

- ✓ المعالجة المسبقة للبيانات.
- ✓ تصنيف العمليات بالاعتماد على الشبكة العصبونية التفاضلية.
- ✓ تقييم أداء النموذج.
- ✓ إجراء التصنيف باستخدام خوارزميات ذكاء صناعي أخرى.
- ✓ أعطت الشبكة العصبونية التفاضلية دقة أفضل من الخوارزميات الأخرى التي تمت المقارنة

معها.

الكلمات المفتاحية: الذكاء الصناعي، الشبكات العصبونية التفاضلية، بطاقات الائتمان، عمليات الاحتيال، التعلم العميق.

Designing a CNN for detecting frude in credit card transactions

Noura kways*

(Received 25/10/2023.Accepted 4/12/2023)

□ABSTRACT □

Financial services are experiencing rapid advancements in the modern era, with advanced technologies being adopted to facilitate and enhance financial transactions. Credit cards are one of the key tools in this context, providing an easy and convenient means of payment and purchase both online and in retail stores.

However, the increased usage of credit cards is accompanied by a rise in financial fraud incidents. Fraud involves obtaining credit card information or illegally exploiting customer data.

To combat this growing phenomenon, credit card companies and financial service providers rely on artificial intelligence (AI) technology to protect customers and achieve higher security in financial transactions. AI is a powerful tool in fraud detection and prevention

The aim of this research is to develop a technique capable of detecting credit card fraud with the highest possible accuracy and a high safety ratio by relying on training a set of artificial intelligence algorithms to classify fraudulent transactions.

The research presents a deep learning model capable of distinguishing between fraudulent and real transactions performed using credit cards. This methodology relied on:

- ✓ Preprocessing of data.
- ✓ Classifying transactions using a convolutional neural network.
- ✓ Evaluating the model's performance.
- ✓ Using other artificial intelligence algorithms for classification.
- ✓ The convolutional neural network gave better accuracy than the other algorithms with which it was compared

Keywords: Artificial intelligence, Convolutional neural networks, Credit cards, Fraudulent activities, Deep learning.

*Postgraduate Department of Communication Technology, Faculty of Information and Communication Technology, University of Tartous, Syria.

نموذج تعلم عميق للمساهمة في كشف الاحتيال على بطاقات الائتمان

١- المقدمة

سمحت خدمة الدفع الإلكتروني بإجراء العمليات المالية (دفع / تحويل / سداد) بشكل إلكتروني بعيداً عن النقود الورقية، إذ تحول الأموال اعتماداً على شفرات رقمية سرية لا يعرفها سوى العميل والجهة التي يتعامل معها.

أدى تعدد وسائل الدفع الإلكتروني، إلى ظهور أنواع مختلفة من بطاقات الدفع بحيث تتفاوت هذه البطاقات في الخصائص والشروط، بعض هذه البطاقات يعمل بناء على الرصيد المتاح للعميل كبطاقة الخصم المباشر Debit Card التي تُتيح للعميل المجال لسحب النقود من المبلغ المودع في حسابه بخلاف بطاقة السحب العادي ATM card التي يُخصم فيها مبلغ الشراء من حساب البنك مباشرة.

تعد بطاقات الائتمان Credit Card نوع من أنواع القروض قصيرة الأمد بحيث تتضمن مبلغ من المال يمكن استخدامه لإجراء عمليات الشراء، ويُخفّض الرصيد المتاح بقدر استخدام البطاقة، ومن ثم يقوم العميل بسداد ما تم إنفاقه من حد الائتمان إلى البنك أو شركة الائتمان فعندما تتم عملية الشراء عبر الإنترنت فإن تفاصيل البطاقة تُرسل إلى البنك، فيحصل عندها البنك على تفويض من شبكة بطاقة الائتمان لإتمام معالجة العملية، بعدها يجب على جهة إصدار بطاقة الائتمان التحقق من بيانات العميل والموافقة على المعاملة أو رفضها.

في حال تمت الموافقة على العملية، يُسد المبلغ إلى البنك المطلوب ويُخفّض الرصيد المتاح بمقدار مبلغ العملية. في نهاية الدورة المالية يتم تلقي بيان يوضح جميع المعاملات التي تمت خلال الشهر، ويوضح الرصيد السابق والمتاح والمبلغ الواجب دفعه وتاريخ الاستحقاق.

تكمن المشكلة في هذا النوع من البطاقات بإمكانية الاحتيال عليها من خلال [٣]:

- الاحتيال دون وجود البطاقة: هو النوع الأكثر شيوعاً، يحدث عندما تُسرق معلومات صاحب البطاقة وتُستخدم استخداماً غير قانوني دون الحاجة إلى وجود البطاقة فعلياً. يحدث هذا النوع عادةً على الإنترنت، وقد يكون نتيجة ما يسمى رسائل التصيد الاحتيالي الإلكتروني التي تُرسل من محتالين ينتحلون شخصية مؤسسات موثوقة بغرض سرقة التفاصيل الشخصية أو المالية عبر رابط. الالكتروني

- احتيال بوجود البطاقة: يأخذ غالباً شكل سرقة معلومات البطاقة باستخدام جهاز إلكتروني للاستنساخ الاحتيالي للبيانات بحيث يمرر البائع البطاقة عبر جهاز يخزن معلومات البطاقة، وتستخدم هذه المعلومات في أية عملية شراء، إذ تُحمّل التكلفة على حساب صاحب البطاقة.

أدت الزيادة المستمرة لعمليات الاحتيال على بطاقات الائتمان إلى مشكلة كبيرة في القطاع المالي وفقاً لذلك استخدمت أساليب وتقنيات التعلم الآلي للكشف عن عمليات الاحتيال ولكنها لم تتمكن من تقديم نتائج بالكفاءة المطلوبة لذلك تم الاعتماد على طرائق التعلم العميق مثل الشبكة العصبونية الالتفافية (CNN) و الشبكات العصبونية ذات الذاكرة طويلة قصيرة الأمد (LSTM) والتي أظهرت نتائج فعالة لكشف الاحتيال نظراً لارتباط طرائق التعلم العميق بمعالجة الصور و البرمجة اللغوية العصبية إذ أظهرت الدراسة أن استخدام LSTM مكون من ٥٠ كتلة يعطي دقة بنسبة ٨٤.٨٥٪ بالنسبة للكشف عن الاحتيال في بطاقات الائتمان والتي تعد أفضل بكثير من الخوارزميات التقليدية [١].

كما تم استخدام التقنيات التحليلية التنبؤية (Predictive Analytic Technologies) للكشف عن عمليات الاحتيال على بطاقات الائتمان، إذ استخدم PAT نماذج التعلم الآلي والنماذج الإحصائية لإنشاء التنبؤات المستقبلية والتي تتضمن خمس مراحل وهي: تحديد المشكلة، الحصول على البيانات وإعدادها، تحليل البيانات وصياغة النموذج، نشر النموذج التنبؤي واختبار أداء النموذج.

اعتمد PAT على مخططات العلامات الحمراء لإجراء تنبؤات لعمليات الشراء غير الشائعة التي يقوم بها حامل البطاقة، وعمليات الشراء المفاجئة على بطاقة الائتمان نفسها، والشحنات المتعددة للبطاقات إلى عنوان واحد، والمعاملات المتعددة على البطاقة في وقت قصير، بالإضافة إلى تحديد الموقع الجغرافي للمعاملة مقارنة بالموقع المسجل لحاملي البطاقات واستخدام عنوان IP واحد لبطاقات ائتمان متعددة لكن تكمن المشكلة في استخدام PAT في قيود النموذج المتمثلة في تكلفة التنفيذ والتعقيد، ومحدودية التدريب وكفاءة التعلم وعدم القدرة على التكيف مع أساليب الاحتيال [٢].

نظراً لأن بعض عمليات الاحتيال تتم مرة واحدة من خلال تقنية الإنترنت ثم يتم استخدام تقنيات أخرى دعت الحاجة إلى وجود أنظمة كشف احتيال تعمل على اكتشاف المعاملات عبر الإنترنت وفقاً لذلك تم إنشاء نموذج بالاعتماد على التعلم العميق والتشفير التلقائي (Auto-Encoder) والذي يعد خوارزمية تعلم غير خاضعة للإشراف تقوم بتطبيق النشر العكسي عن طريق تعيين دخل مساوي للخروج وآلة Boltzmann المقيدة (Restricted Boltzmann Machine) المؤلفة من طبقتين، طبقة الإدخال (مرئية) والطبقة المخفية والتي يتم من خلالها إعادة بناء المعاملات العادية للعثور على الحالات الشاذة من الأنماط العادية [٤].

يظهر الجدول (١) مقارنة أساليب الدراسات المرجعية:

الدراسة	الحل المقترح	النتائج
الدراسة [١]	استخدام خوارزمية LSTM	تصنيف بدقة ٨٤.٨٥٪ للكشف عن الاحتيال في بطاقات الائتمان
الدراسة [٢]	استخدام التقنيات التحليلية التنبؤية (Predictive Analytic Technologies)	تكلفة التنفيذ والتعقيد، ومحدودية في التدريب وكفاءة التعلم وعدم القدرة على التكيف مع أساليب الاحتيال
الدراسة [٤]	استخدام التعلم العميق والتشفير التلقائي	التصنيف بدقة أعلى من الخوارزميات التقليدية

تم البحث ضمن هذا العمل على تحسين نتائج الدراسة المرجعية [٤] من خلال استخدام الشبكة العصبونية الالتفافية مع تقنية التسرب وتقنية Batch normalization

٢- أهمية البحث وأهدافه

تسمح بطاقات الائتمان بإجراء عمليات الشراء الكبيرة التي تتطلب مبالغ ضخمة والتي تحتاج إلى وقت طويل لإدخالها ومن ثم تقسيم هذه التكلفة على مدى زمني طويل بالإضافة إلى توفير كشف حساب شهري بالمعاملات المالية، مما يجعل تتبع النفقات أسهل بكثير.

ولكن أدى الازدياد الكبير لعمليات الاحتيال على بطاقات الائتمان في السنوات الأخيرة الى تكليف العملاء والشركات مئات الملايين من الدولارات سنويًا وتتمثل عمليات الاحتيال من خلال الحصول على أرقام بطاقة الائتمان أو الخصم الخاصة بالعميل من خلال مواقع الويب غير المحمية أو من خلال سرقة معلومات هوية العميل.

وتكمن أهمية قدرة شركات بطاقات الائتمان على التعرف على معاملات البطاقات الصادرة من خلالها حتى لا يتم تحصيل رسوم من العملاء مقابل العناصر التي لم يتم شراؤها. وبناءً على ما تقدم فإن الهدف من البحث هو بناء نموذج تعلم عميق قادر على التعرف على العمليات الاحتيالية بأعلى دقة ممكنة.

٣- طرائق البحث ومواده

يساهم هذا البحث في إيجاد نموذج تعلم عميق يعمل على كشف الاحتيال الحاصل على بطاقات الائتمان وبكفاءة عالية وفق التالي:

□ إجراء عمليات المعالجة المسبقة على قاعدة البيانات وموازنة مجموعة البيانات لجعل الفئات متشابهة.

□ تصنيف العمليات بالاعتماد على الشبكة العصبونية الالتفافية. CNN

□ اختبار وتقييم أداء النموذج.

نُفذ هذا المنهج باستخدام المحاكاة الحاسوبية من خلال:

□ بيئة Anaconda ومن خلاله تم العمل على برنامج Jupyter Notebook لتوفيره

المكتبات الداعمة لمثل هذه الأبحاث.

□ الاعتماد برمجياً على لغة البايثون Python الإصدار ٣.٦ بسبب احتوائها على المكتبات

الداعمة للبحث

□ مجموعة بيانات تتضمن العمليات التي تم إجراؤها بواسطة بطاقات الائتمان في أيلول

٢٠١٣.

٤- الدراسات المرجعية (Literature Review)

٤-١ التعلم العميق

التعلم العميق هو حقل فرعي من تعلم الآلة (machine learning) والذي يهتم بالخوارزميات ويعتمد على بنية ووظائف الدماغ من خلال الاعتماد على التعلم متعدد المستويات المقابل لبنية هرمية من السمات أو المفاهيم عالية المستوى يتم تعريفها بالاعتماد على المفاهيم ذات المستوى الأدنى ويمكن وضع مفاهيم أساسية للتعلم العميق وهي:

١. وجود نماذج مؤلفة من عدة مستويات من المعالجة غير الخطية للمعلومات.

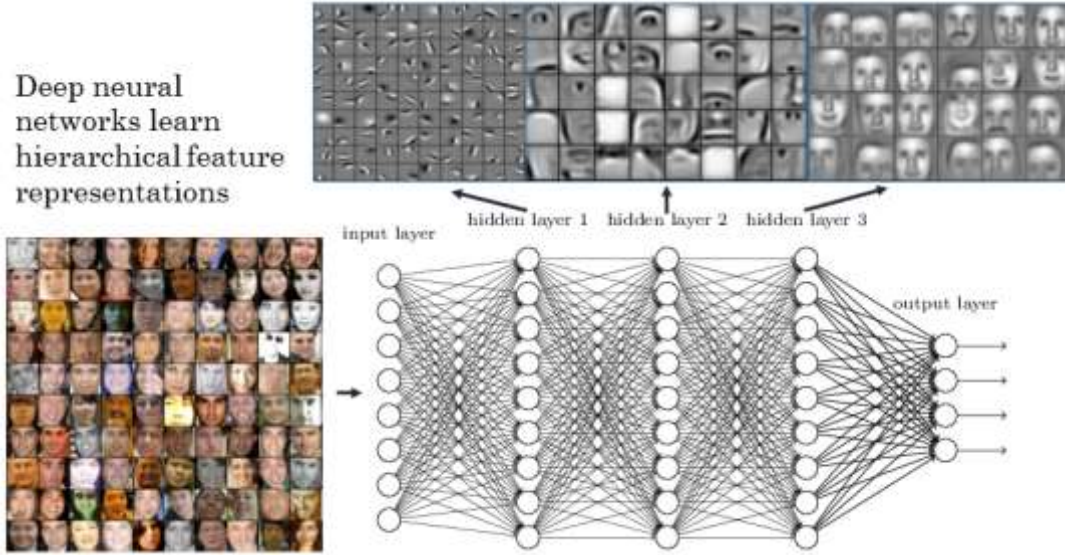
٢. طرائق تعلم السمات (بإشراف أو بدون إشراف) أكثر تجريباً في المستويات الأعلى.

يقع التعلّم العميق في المنطقة المشتركة ما بين عدّة مجالات وهي بحوث الشبكات العصبونية والذكاء الصناعي والنمذجة الرسومية والأمثلة والتعرّف على الأنماط، وهناك ثلاثة أسباب وراء انتشار التعلّم العميق وشيوعه وهي:

١. القدرات المتزايدة لمعالجة الرقاقات، كوحدات المعالجة الرسومية عامة الأغراض.
٢. الحجم المتزايد للمعطيات المستخدمة في التدريب.
٣. التقدّم الحديث في تعلّم الآلة ومعالجة الإشارة والمعلومات.

وقد سمحت هذه التطورات الأخيرة لوسائل التعلّم العميق باستعمال التتابع المعقّدة المركّبة غير الخطية من أجل تعلّم سمات الخواص الموزّعة والهرمية، وتحقيق الاستعمال الفعّال للمعطيات المصنّفة وغير المصنّفة. ظهر التعلّم العميق من المجال البحثي الخاص بشبكات الذكاء الصناعي العصبونية، وتعدّ الشبكات العصبونية ذات التغذية المسبقة والمجهّزة بعدّة طبقات مخفية وتدعى الشبكات العصبونية العميقة مثلاً جيداً عن نماذج البنى العميقة.

تعد خوارزمية الانتشار الخلفي (BP Back-propagation)، خوارزمية معروفة لتعلّم معاملات هذه الشبكات، ولكن الانتشار الخلفي لوحده لم يعمل بشكلٍ جيّدٍ مع شبكات التعلّم التي تملك عدداً كبيراً من الطبقات المخفية، وقد كانت الصعوبة تزداد كلما ازداد عمق هذه الشبكات.



الشكل (١) الشبكة العصبونية العميقة

يمثل الشكل (١) كيف تقوم هذه الشبكات العصبونية العميقة بتعلّم التمثيلات الهرمية للسمات. تم استخدام الطبقات المخفية ضمن العديد من العصبونات ضمن الشبكة العصبونية العميقة مما أدى الى تحسّن قوة نمذجة هذه الشبكات والتي بدأت بتشكيل العديد من النماذج المثالية، فحتى وإن تعطلّ مسار متحوّلٍ ما، ستتابع الشبكة العصبونية العميقة الناتجة أداء عملها بشكلٍ جيّدٍ [5].

Tensorflow 2-4

مكتبة مفتوحة المصدر تستخدم لغة بايثون Python للحساب الرقمي والذي يجعل تعلم الآلة أسرع وأسهل يتيح مكتبة TensorFlow إنشاء رسومات بيانية لتدفق البيانات وهي عبارة عن هياكل تصف كيفية تحرك البيانات عبر الرسم

البياني، أو سلسلة من عقد المعالجة وتمثل كل عقدة في الرسم البياني عملية حسابية، ويكون كل اتصال أو حافة بين العقد عبارة عن مصفوفة بيانات متعددة الأبعاد.

يوفر TensorFlow طرقاً سهلة للتعبير عن كيفية الجمع بين التجريدات عالية المستوى، العقد والتوترات في TensorFlow هي كائنات Python، وتطبيقات TensorFlow هي نفسها تطبيقات Python. ومع ذلك، لا يتم تنفيذ العمليات الفعلية الرياضية في Python وإنما تتم كتابة مكتبات التحويلات التي تتوفر من خلال TensorFlow كثنائيات ++C عالية الأداء، لتقوم Python فقط بتوجيه حركة المرور بين القطع، وتوفر اختصارات برمجة عالية المستوى لربطها ببعضها البعض. الفائدة الأكبر التي يوفرها TensorFlow لتنمية تعلم الآلة هي التجريد. وبدلاً من التعامل مع التفاصيل الدقيقة لتنفيذ الخوارزميات، أو اكتشاف الطرق المناسبة لمخرجات دالة واحدة إلى مدخلات أخرى، يمكن للمطور التركيز على المنطق العام للتطبيق. ويعتني TensorFlow بالتفاصيل الموجودة. [٧]

Keras 3-4

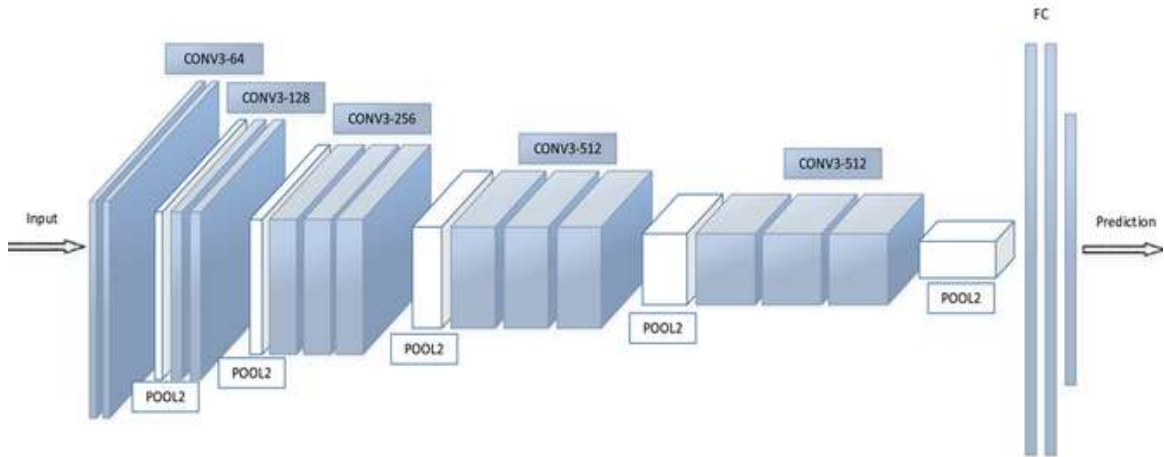
هي مكتبة شبكة عصبونية مفتوحة المصدر مكتوبة بلغة Python تعمل فوق Theano أو TensorFlow وهي مصممة لتكون معيارية وسريعة وسهلة الاستخدام لا تتعامل Keras مع الحسابات ذات المستوى المنخفض وإنما تقوم باستخدام مكتبة Backend.

Keras عبارة عن غلاف عالي المستوى لواجهة برمجة التطبيقات Application API (Programming Interface)

منخفضة المستوى وهو قادر على العمل فوق TensorFlow أو CNTK أو Theano. تتعامل Keras High-Level API مع الطريقة التي نضع بها النماذج أو نحدد الطبقات أو من خلال وضع نماذج متعددة للإدخال والإخراج وضمن هذا المستوى يقوم Keras أيضاً بتجميع نموذجنا مع دوال loss و optimizer واختيار دالة تدريب مناسبة [8]

4- ٤ الشبكات العصبونية الالتفافية CNN

تعمل هذه الشبكة الموضحة في الشكل (٢) مع الصور ضمن مجال التعلم العميق [٦]، بحيث يكون دخل الشبكة العصبونية هي عبارة عن صور أو بشكل أكثر تحديد مصفوفة ثلاثية الأبعاد، وتتألف هذه الشبكات من سلسلة من الطبقات والتي تتعلم استخراج السمات المميزة من أي صورة.



الشكل (2) الشبكة العصبونية الالتفافية

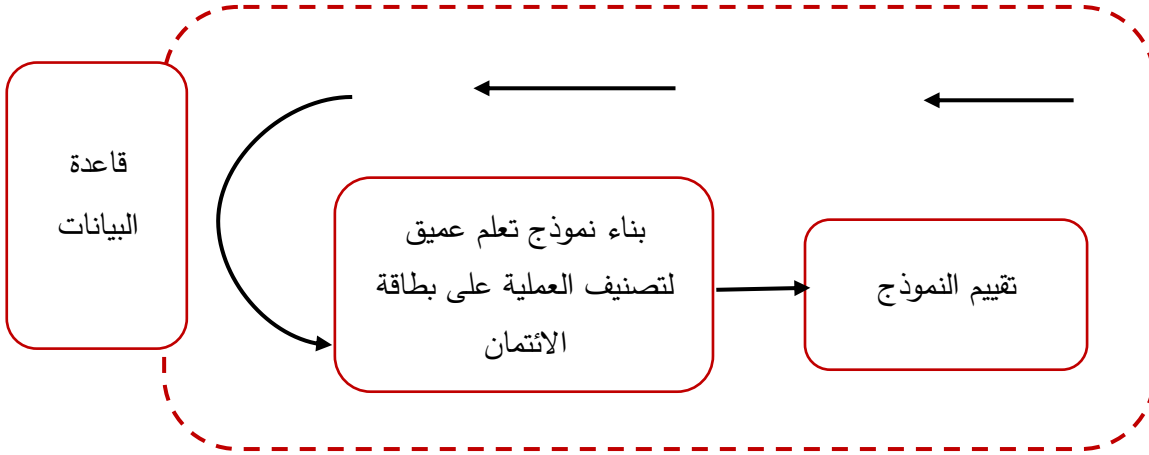
- طبقة التلافية ((CONV): تعتبر الأساس في هذا النوع من الشبكات العصبية ، والتي تقوم بتطبيق سلسلة من مرشحات الصور (filters) المختلفة على الصورة المدخلة، وهذه المرشحات تستخرج سمات مختلفة من الصورة مثل حواف الأجسام والزوايا والتدرجات اللونية وبينما تتدرب شبكات الطبقة الالتفافية العصبية فهي تقوم بتحديث الأوزان ضمن هذه الطبقة باستخدام الانتشار العكسي (backpropagation) وهذه الأوزان بدورها تحدد نوعية مرشح الصورة، والنتيجة النهائية هي مُصنّف (classifier) يتألف من العديد من الطبقات الالتفافية والتي بدورها تعلمت كيفية ترشيح الصورة لاستخراج السمات (features) المهمة منها وتشير (CONV3- CONV 128.....-364، 012) إلى عدد الطبقات الالتفافية والمرشحات المستخدمة ضمنها.
- طبقة التجميع (pool): تعد هذه الطبقات اختيارية في تصميم الشبكة، وفي حال وجودها سيكون موقعها بعد كل طبقة من الطبقات الالتفافية وتهدف إلى تخفيض عدد العينات أو العصبونات حيث ستقوم باختصار كل مجموعة من عصبونات الدخل بحجم معين إلى عصبون واحد، ويحدد هذا الحجم ضمن تصميم الشبكة وتكون قيمته المثلى 2×2 لأن تكبيرها قد يؤدي إلى ضياع في المعلومات ويتم التخفيض بعدة طرق منها:
 - القيمة العليا المشتركة: (max pooling) تأخذ القيمة الأعلى بينها.
 - المعدل المشترك: (average pooling) تأخذ معدل جميع القيم
- طبقة الاتصال الكامل (FC):

بعد عدة طبقات من النوعين السابقين تأتي هذه الطبقات لتربط كل عصبونات الطبقة السابقة (مهما كان نوعها) وتجعلها دخل لكل عصبون فيها كما في الشبكات العصبونية العادية، لا يشترط أن تكون بعدد معين ولكن غالباً يوجد منها طبقتان متتاليتان كالتبقيات الأخيرة في الشبكة إذ لا يمكن أن تأتي قبل طبقة من النوع الالتفافي.

5 - المناقشة

نُفذ البحث بتطبيق سلسلة من الخطوات لكشف الاحتيال على بطاقة الائتمان ويمثل المخطط التالي المبين بالشكل

(٣) المنهجية المقترحة



الشكل (3) خطوات العمل

سيتم إجراء المعالجة المسبقة لقاعدة البيانات وبناء الشبكة العصبونية الالتفافية التي ستقوم بتصنيف المميزات الخاصة لكل بطاقة ائتمان موجودة ضمن قاعدة البيانات بشكل يؤدي هذا التصنيف الى أعلى دقة ممكنة.

١-٥ المعالجة المسبقة لقاعدة مجموعة البيانات:

تم العمل على قاعدة البيانات التي تم تنزيلها من موقع Kaggle من خلال الرابط <https://www.kaggle.com/mlg-ulb/creditcardfraud> وتم استيراد:

- مكتبة Numpy والتي تعمل على معالجة البيانات وفق:
- تحويل البيانات إلى النطاق [٠,١] و [-١,١]، وذلك لتسهيل عمليات المعالجة اللاحقة.
- استبدال القيم المفقودة بقيمة محددة أو حذف الصفوف أو الأعمدة التي تحتوي على قيم مفقودة.
- تقسيم البيانات إلى جزئيات أصغر ومعالجتها بشكل مستقل، وذلك لتحسين سرعة وكفاءة المعالجة
- مكتبة pandas والتي تعمل على تحليل البيانات من خلال:
- قراءة البيانات من ملف CSV.
- التعرف على بنية البيانات: عدد الأعمدة والصفوف وأنواع البيانات المختلفة المتوفرة في الجدول.
- تصفية البيانات واستبعاد الأعمدة أو الصفوف التي لا تحتوي على بيانات مهمة.
- تحويل البيانات إلى شكل آخر يمكن فهمه والتعامل معه، مثل تحويل البيانات إلى جداول مفصلة أو مخططات بيانية أو قوائم.
- إجراء التحليل الإحصائي على البيانات، مثل حساب المتوسط والانحراف المعياري والحد الأدنى والحد الأقصى وتحليل الترابط بين البيانات.

Time	V1	V2	V3	V4	V5	V6	V7	V8	V9	...	V21	V22	V23	V24		
0	0.0	-1.359807	-0.072781	2.536347	1.378155	-0.338321	0.462388	0.239599	0.098698	0.363787	...	-0.018307	0.277838	-0.110474	0.066928	0.
1	0.0	1.191857	0.266151	0.166480	0.448154	0.060018	-0.082361	-0.078803	0.085102	-0.255425	...	-0.225775	-0.638672	0.101288	-0.339846	0.
2	1.0	-1.358354	-1.340163	1.773209	0.379780	-0.503198	1.800499	0.791461	0.247676	-1.514654	...	0.247998	0.771679	0.909412	-0.689281	-0.
3	1.0	-0.966272	-0.185226	1.792993	-0.863291	-0.010309	1.247203	0.237609	0.377436	-1.387024	...	-0.108300	0.005274	-0.190321	-1.175575	0.
4	2.0	-1.158233	0.877737	1.548718	0.403034	-0.407193	0.095921	0.592941	-0.270533	0.817739	...	-0.009431	0.798278	-0.137458	0.141267	-0.

الشكل (4) قاعدة البيانات

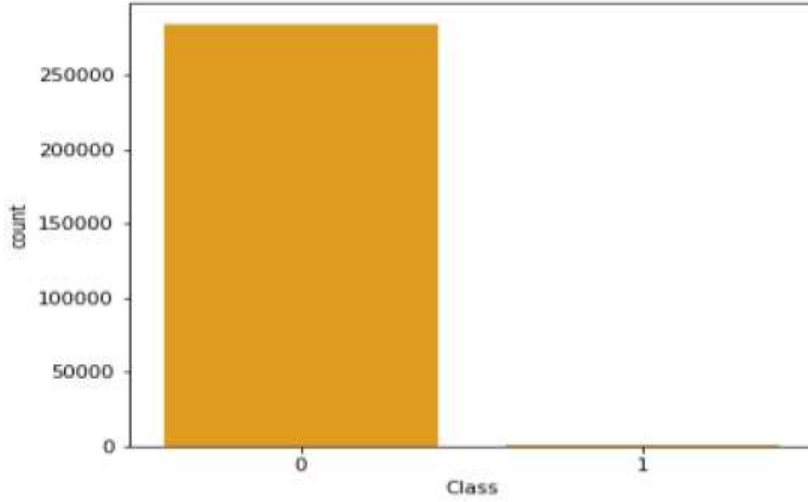
تتضمن قاعدة البيانات الموضحة الشكل (٤):

- عدد العمليات الاحتمالية ٤٩٢ والعمليات الحقيقية ٢٨٤٣١٥
- الميزات V-٢٨٧١ وهي تمثيلات رقمية ناتجة عن تحويل المكون الرئيسي (PCA) تشمل الميزات V-٢٨٧١ مجموعة من المتغيرات الرقمية التي تمثل نتائج تحويل PCA للبيانات الأصلية، وتشمل عادة معلومات مثل الرصيد الحالي للبطاقة وعدد المعاملات والمتوسطات والانحرافات المعيارية والتواريخ

والأوقات والتي يمكن استخدامها لتحديد الأنماط والسلوكيات غير المعتادة في استخدام البطاقات الائتمانية والتي يمكن أن تشير إلى وجود نشاط احتيالي.

● "Amount" قيمة العملية بالدولار .

● متغير " Time " هو مقدار الوقت المقتضي من وقت حدوث العملية الأولى (العملية الأولى قد تم تنفيذها في وقت محدد، ومن ثم تم قياس المدة الزمنية المستغرقة في تنفيذ العمليات التالية بعد العملية الأولى، وجمع الناتج مع الوقت الذي استغرق في تنفيذ العملية الأولى للحصول على إجمالي الوقت المستغرق في تنفيذ العملية بالكامل)
● الاحتيال = 1، عدم الاحتيال = 0.



الشكل (5) تمثيل البيانات

يوضح الشكل (5) التمثيل المرئي للبيانات والذي يظهر كيفية توزيع البيانات بين العمليات الاحتمالية والحقيقية ويشير الشكل الى عدم توازن البيانات فإذا لم تكن معظم العمليات احتمالية، سيصبح من الصعب تصنيف بعض العمليات على أنها حقيقية أم لا لذلك تم استخدام عمود الفئة class لحساب عدد العمليات الاحتمالية ومن ثم حساب النسبة المئوية لها.

تشير المعاملات الإحصائية أيضا الى عدم التوازن بين العمليات الاحتمالية والحقيقية من خلال تحليل النسب وتحليل البيانات الزمنية وفق الشكل (6) والذي يتضمن عدد الحالات (count) والمتوسط (mean) والانحراف المعياري (std)

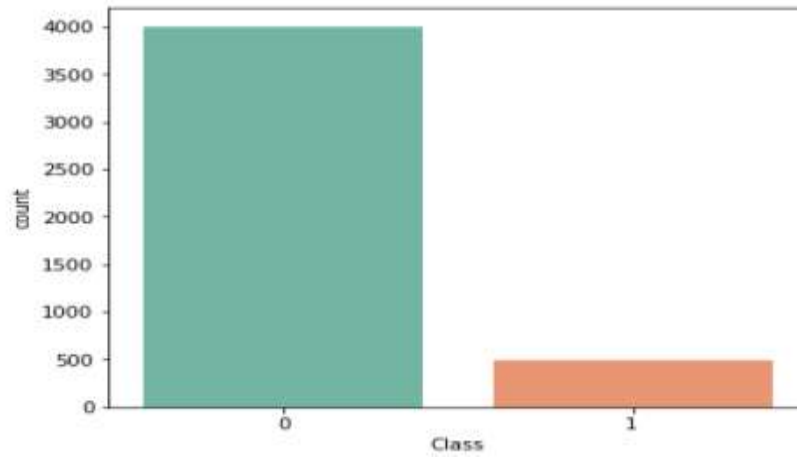
```

False Detection Cases
-----
count      492.000000
mean       122.211321
std        256.683288
min         0.000000
25%         1.000000
50%         9.250000
75%        105.890000
max        2125.870000
Name: Amount, dtype: float64

True Detection Cases
-----
count      284315.000000
mean         88.291022
std         250.105092
min          0.000000
25%          5.650000
50%         22.000000
75%         77.050000
max        25691.160000
Name: Amount, dtype: float64
    
```

الشكل (6) القيم الإحصائية

يظهر الشكل رقم (٧) تقنية موازنة البيانات في تحليل الاحتمال على بطاقات الائتمان من خلال اقتطاع نطاقات محددة القيمة، وهذا يعني استخدام عينة من البيانات بدلاً من استخدام جميع البيانات المتاحة. ويتم اختيار هذه النطاقات بحيث تشكل نسبة محددة من البيانات الإيجابية (الحقيقية) والبيانات السلبية (الاحتمالية) وبالتالي أصبحت الفئات قريبة من بعضها



الشكل (7) موازنة البيانات

٢-٥ نموذج التعلم العميق:

خطوات العمل:

- تفعيل بيئة للتعلم العميق من خلال نافذة الأوامر Anaconda prompt حيث سيتم تثبيت هذا النموذج ضمنها
- تثبيت مكتبة Tensorflow

● تثبيت مكتبة Keras.

● بناء الشبكة العصبونية التفاضلية (CNN):

١-٢-٥ الشبكة العصبونية التفاضلية:

تم تقسيم البيانات الى ٨٠% تدريب و ٢٠% اختبار واستخدام محسن التصنيف adam وهو عبارة عن خوارزمية تقوم بتسريع أداء النموذج أي تجعله يتقارب إلى القيم الصغرى الشاملة ل Cost Function بشكل أسرع وأدق وفق الخطوات:

١. تعيين قيم البداية: معدل التعلم (٠.٠٠١) وعوامل الانحراف المعياري (١beta و ٢beta) ومستوى الدقة (epsilon).

٢. حساب المشتقات الجزئية لدالة التكلفة (cost function) بالنسبة لكل معامل (weights) في الشبكة العصبونية.

٣. حساب متوسط (momentum) لكل معامل باستخدام عامل الانحراف المعياري ١beta، وتحديث قيم المتوسطات.

٤. حساب متوسط المربعات (squared momentum) لكل معامل باستخدام عامل الانحراف المعياري ٢beta، وتحديث قيم المتوسطات.

٥. تعديل معدل التعلم (learning rate) باستخدام قيم المتوسطات المحسوبة في الخطوات السابقتين.

٦. تحديث قيم المعاملات (weights) باستخدام المتوسطات المحسوبة في الخطوات السابقتين ومعدل التعلم المعدل في الخطوة السابقة.

٧. تكرار الخطوات من ٢ إلى ٦ حتى تتحقق شروط الانتهاء من التدريب (مثل عدد الدورات التدريبية المطلوبة أو تحقيق مستوى الدقة المطلوب).

يتم تنفيذ هذه الخطوات في كل دورة تدريبية، وتؤدي إلى تحسين معدل التعلم وسرعة التدريب وجودة التصنيف في الشبكات العصبونية الاصطناعية مع دالة خسارة Binary Crossentropy وتم تدريب النموذج لمدة ٢٠ مرة.

يظهر الشكل (٨) ملخص لنموذج التعلم العميق المستخدم ضمن بحثنا حيث تضمن:

- الطبقة الأولى: هي طبقة التفاضلية كاملة الاتصال مع استخدام ٣٢ مرشح بأبعاد ٢*٢ مع تابع تفعيل Relu .
- الطبقة الثانية: هي طبقة التفاضلية مخفية كاملة الاتصال وتم استخدام تقنية Batch normalization وهي أسلوب للتدريب يتم إجراؤه بين طبقات الشبكة العصبونية بدلاً من البيانات الأولية بحيث يتم إجراء التدريب على بيانات صغيرة بدلاً من مجموعة البيانات الكاملة حيث أنه يعمل على تسريع التدريب واستخدام معدلات التعلم الأعلى، مما يجعل التعلم أسهل ويقلل عدد مرات التدريب مع ٦٤ مرشح بأبعاد ٢*٢ مع تابع تفعيل Relu مع فرصة تسرب بمقدار ٠.٢ (الهدف من استخدام فرصة التسرب هو استبعاد بعض العصبونات في الطبقات المخفية بهدف تجاوز حدوث حالة (over-fitting)

● طبقة التنسوية (Fattening): تم استخدام فرسنتي تسرب ٠.٤ و ٠.٥ مع ٦٤ مرشح وتابع تفعيل Relu .

● الطبقة الأخيرة: طبقة الخرج للشبكة العصبونية كاملة الاتصال مع تابع تفعيل Sigmoid الذي يعطي الناتج ضمن نطاق [٠، ١] ويناسب دالة الخسارة.

يجب ملاحظة انه تم استخدام تابع Sigmoid بدلاً من التابع hardlim، لأنه في حالة مسألة الاحتيال، يتم تصنيف المدخلات (inputs) إما كاحتيال (fraud) أو كعملية حقيقية (legitimate)، ولذلك تعتبر مسألة تصنيف

ثنائية. وبالتالي، يكون الخرج المطلوب من الشبكة هو رقم يتوافق مع الاحتمالية لكل من الحالتين (fraud, legitimate)، وليس قرار ثنائي صريح. وبالتالي، يمكن استخدام تابع Sigmoid كوظيفة تنشيط للخلايا العصبية النهائية في الشبكة العصبية الاصطناعية، حيث يتم تحويل الخرج إلى نطاق ٠-١، مما يتيح تفسيره كاحتمالية للحالة الإيجابية (fraud) بالنسبة إلى الحالة السلبية (legitimate). بالمقارنة، فإن التابع hardlim هو تابع تفاعلي ثنائي القيمة (binary) يستخدم لحالات التصنيف الثنائية فقط، ولا يمكن استخدامه لتحويل الخرج إلى احتمالية أو قيمة مستمرة. لذلك، يمكن استخدام التابع hardlim في حالات مسائل التصنيف الثنائي فقط، وليس لحالات الاحتمالات والتصنيف الثنائي المستمر.

Model: "sequential"

Layer (type)	Output Shape	Param #
conv1d (Conv1D)	(None, 29, 32)	96
dropout (Dropout)	(None, 29, 32)	0
batch_normalization (Batch Normalization)	(None, 29, 32)	128
conv1d_1 (Conv1D)	(None, 28, 64)	4160
dropout_1 (Dropout)	(None, 28, 64)	0
flatten (Flatten)	(None, 1792)	0
dropout_2 (Dropout)	(None, 1792)	0
dense (Dense)	(None, 64)	114752
dropout_3 (Dropout)	(None, 64)	0
dense_1 (Dense)	(None, 1)	65

=====
 Total params: 119,201
 Trainable params: 119,137
 Non-trainable params: 64

الشكل (8) نموذج التعلم العميق

٦- النتائج:

٦-١ مصفوفة الارتباك

تعد مقياس الأداء لمشكلة التصنيف بحيث تعد:

- التصنيفات الصحيحة هي عبارة عن الإيجابيات الحقيقية TP (التنبؤ بالاحتيال وهو قد حدث فعلاً) أو السلبيات الحقيقية TN (التنبؤ بعدم وجود احتيال)
- التصنيفات الخاطئة هي عبارة عن الإيجابيات الكاذبة FP (التنبؤ بالاحتيال وذلك غير صحيح) أو السلبيات الخاطئة FN (التنبؤ بعدم وجود احتيال وهو غير صحيح)

وبالاعتماد على هذه القيم يتم استخلاص الحساسية والنوعية والدقة.

		Target class	
Output class	3	TP الإيجابي الحقيقي (791)	FP الإيجابي الكاذب (3)
	9	FN السلبي الكاذب (9)	TN السلبي الحقيقي (96)

الشكل (٩) قيم مصفوفة الارتباك

الحساسية (Sensitivity): هي القدرة على كشف حالات الاحتيال عندما يكون الاحتيال حدث فعلاً [٩] وتحسب

الحساسية وفق العلاقة:

$$\text{Sensitivity} = \frac{TP}{TP+FN} * 100 = \frac{791}{791+9} * 100 = 98.875$$

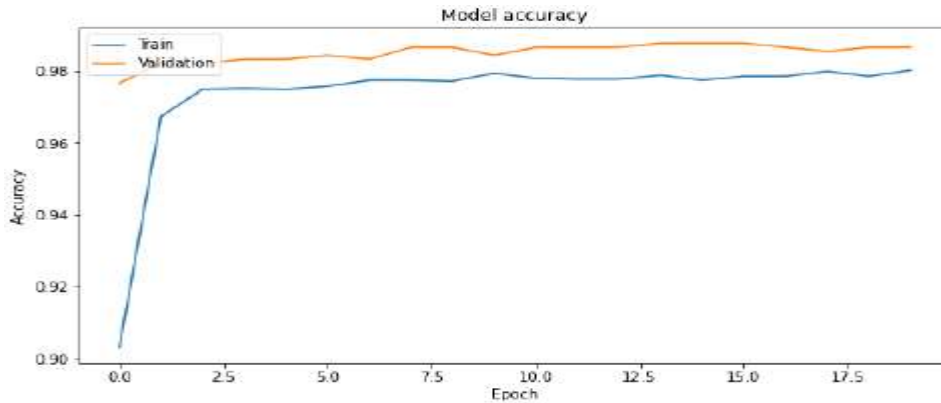
النوعية (Specificity): هي القدرة على تحديد العمليات الحقيقية [٩] وتحسب النوعية وفق العلاقة:

$$\text{Specificity} = \frac{TN}{TN+FP} * 100 = \frac{96}{96+3} * 100 = 96.96$$

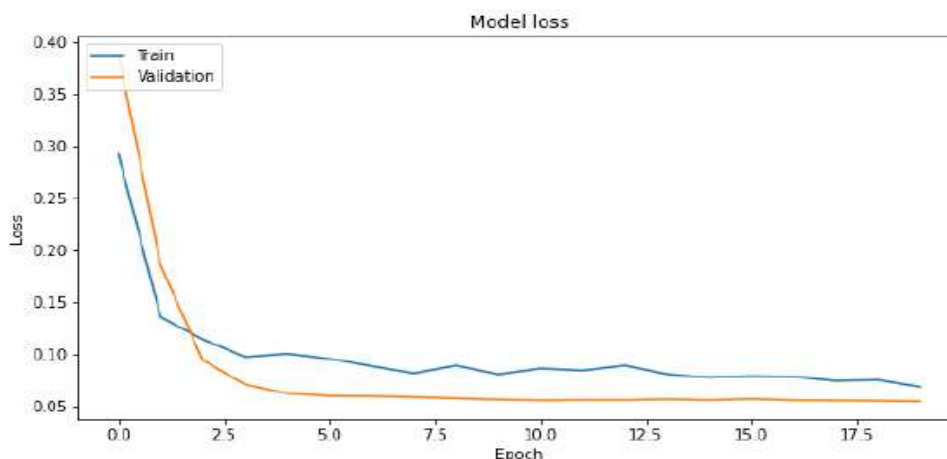
تحسب الدقة [٩] وفق العلاقة:

$$\text{Accuracy} = \frac{TP+TN}{TP+FN+TN+FP} * 100 = \frac{791+96}{791+9+96+3} * 100 = 98.66$$

● التدريب



الشكل (١٠) دقة النموذج



الشكل (١١) أداء النموذج

يمكن الملاحظة من الشكلين (١٠) و(١١) أن الشبكة قد تعلمت عند التكرار ٥ مع أفضل دقة لأداء الشبكة هي 98.66 وأفضل قيمة للتحقق هي ٠.٠٥ هي قريبة جداً من ٠ مما يدل على دقة تصنيف الشبكة وذلك يعود الى استخدام الشبكة العصبونية الالتفافية مع تقنية Batch normalization و هي تقنية تستخدم في تدريب الشبكات العصبية، وتعد واحدة من التقنيات الأساسية التي تساعد في تسريع عملية التدريب وتحسين أداء النموذج وتعتمد تقنية Batch normalization على تحويل الخرج الموجود في كل طبقة من طبقات النموذج، وبشكل أساسي، يتم تحويل الخرج عن طريق حساب المتوسط والانحراف المعياري للخرج في كل دفعة من البيانات التي تمر عبر النموذج، ومن ثم تحويل البيانات وتحويل الخرج باستخدام هذه الإحصائيات. يتم تطبيق Batch normalization على الإخراج الناتج من طبقة معينة، ويتم حساب المتوسط والانحراف المعياري عبر البعد الذي يتم إخرجه من هذه الطبقة. وبعد ذلك، يتم تطبيع الإخراج باستخدام المتوسط والانحراف المعياري، ومن ثم يتم تحويل البيانات باستخدام العوامل المعينة خلال عملية التدريب.

٦-٢ خوارزميات التصنيف الأخرى:

□ تم استخدام قاعدة البيانات نفسها للتصنيف بالاعتماد على خوارزمية الجوار الأقرب والتي تخزن جميع الحالات المتاحة لديها من بيانات التدريب وتقوم بتصنيف الحالات الجديدة حسب أغلبية جيرانها، وذلك عن طريق وظيفة حسابية لقياس المسافة بينهم وضمن البحث قمنا بتحديد ($k=3$) (تم اختياره تجريبياً) وهو عبارة عن عدد عناصر الجوار.

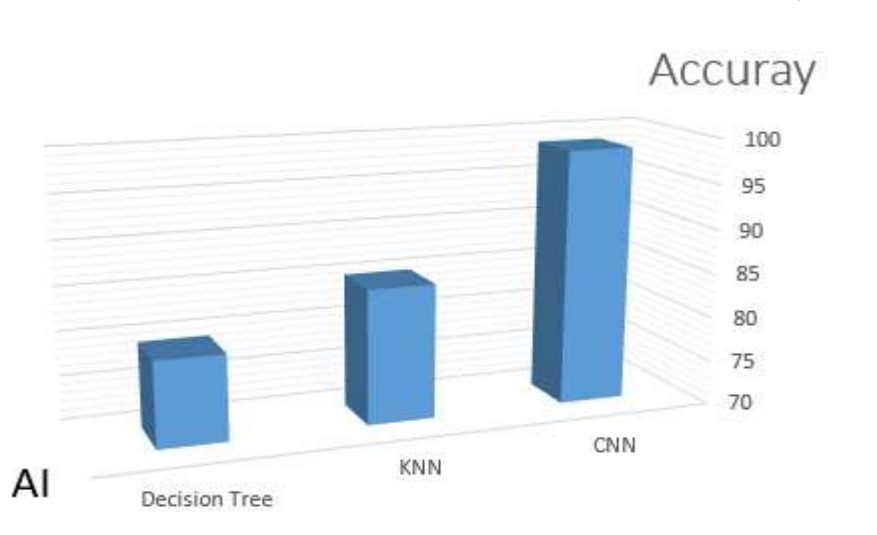
□ شجرة القرار Decision Tree.

ويمكن تلخيص نتائج استخدام هذه الخوارزميات بالجدول (٢):

جدول (٢) ملخص نتائج تطبيق خوارزميات الذكاء الصناعي

AI	Sensitivity	Specificity	Accuracy
CNN	98.875	96.96	98.66
KNN (k=3)	85.5	87.2	85.1
Decision Tree	80.9	78.1	79.6

أعطت الشبكة العصبونية أفضل دقة مقارنة بالخوارزميات التي تم ذكرها وبالاعتماد على قاعدة البيانات نفسها، يظهر المخطط التالي نتيجة المقارنة:



الشكل (١٢) مقارنة دقة خوارزميات الذكاء الصناعي

٧- الاستنتاجات والتوصيات

- ✓ أدى استخدام تقنية التسرب الى التقليل من تعقيد النموذج وبالتالي التقليل من زمن التصنيف .
- ✓ أدى استخدام تقنية Batch normalization الى التقليل عدد مرات تدريب النموذج والى تعلم دقة أعلى .
- ✓ تتمتع الطريقة المقترحة بالسرعة والفعالية والتي تؤدي الى انخفاض العمليات الاحتمالية.
- ✓ أدى استخدام Tensorflow و Keras الي توفير إطار بيانات سهل الاستخدام بالإضافة الى المرونة في بناء الشبكة العصبونية الالتفافية.
- ✓ أدى استخدام الشبكة العصبونية الالتفافية (CNN) الى تصنيف العمليات بأعلى دقة ممكنة وبشكل أفضل من الشبكة خوارزمية الجوار الأقرب وشجرة القرار.

المراجع

- [1] Nguyen, Thanh Thi, et al. "Deep learning methods for credit card fraud detection." *arXiv preprint arXiv:2012.03754* (2020).
- [2] K. T. Hafiz, S. Aghili and P. Zavarisky, "The use of predictive analytics technology to detect credit card fraud in Canada," 2016 11th Iberian Conference on Information Systems and Technologies (CISTI), Las Palmas, 2016, pp. 1-6.
- [3] A. O. Adewumi and A. A. Akinyelu, "A survey of machine-learning and nature-inspired based credit card fraud detection techniques." *International Journal of System Assurance Engineering and Management* 8, no. 2 (2017): 937-953. Pumsirirat, A., & Liu, Y. (2018). [4] [4] *Credit card fraud detection using deep learning based on auto-encoder and restricted boltzmann machine*. *International Journal of advanced computer science and applications*, 9(1).

[5] Gao, L., Yi, X., Liao, Y., Jiang, Z., Yan, Z., & Tang, Z. (2017, November). *A deep learning-based formula detection method for PDF documents*. In 2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR) (Vol. 1, pp. 553-558). IEEE.

[6] Zhang, K., Zhang, Z., Li, Z. and Qiao, Y., 2016. *Joint face detection and alignment using multitask cascaded convolutional networks*. IEEE signal processing letters, 23(10), pp.1499-1503.

[7] Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., ... & Zheng, X. (2016, November). *Tensorflow: a system for large-scale machine learning*. In *Osd* (Vol. 16, No. 2016, pp. 265-283).

[8] Haghghat, E., & Juanes, R. (2021). *SciANN: A Keras/TensorFlow wrapper for scientific computations and physics-informed deep learning using artificial neural networks*. *Computer Methods in Applied Mechanics and Engineering*, 373, 113552.