وزارة التعليم العالي والبحث العلمي
جامعة بابل
كلية التربية للعلوم الصرفة
قسم الرياضيات

Ministry of Higher Education
And Scientific Research
University of Babylon
College of Education for
Pure Sciences
Mathematics Department

التعلم العميق Deep Learning

بحث مقدم الى مجلس كلية التربية للعلوم الصرفة/جامعة بابل

كجزء من متطلبات التخرج لنيل شهادة البكالوريوس

إعداد الطالبة

بنين حسين عبدعون حسون

بإشراف الدكتورة

طفول حسين الخفاجي

1447 A.H

2026 A.D

بِسْمِ اللهِ الرَّحْمنِ الرَّحِيمِ

﴿وَقُلِ اعْمَلُوا فَسَيَرَى اللَّهُ عَمَلَكُمْ وَرَسُولُهُ وَالْمُؤْمِنُونَ﴾

صَدقَ الله العَلي العَظيم

التوبة: ١٠٥

شُكر وتقدير

اود ان اعبر عن خالص امتناني لمشرفتي الدكتورة ..... ولتوجيهاتها والتعليقات البناءة لهذا العمل العملي والذي يعد جزء من متطلبات نيل شهادة البكالوريوس الجامعية لا يسعني سوى ان أقدّم الشكر والعرفان للدكتورة الفاضلة

كما اتقدم الى عمادة كلية التربية ورئاسة قسم الرياضيات بالشكر والعرفان لكل جهودهم التي بذلوها طوال سنوات دراستي واتقدم بالشكر الجزيل لجميع اساتذتي الافاضل الذين بذلوا كل ما في وسعهم ليسلكوا بي طريق العلم والمعرفة.

الإهداء

الى معلم الانسانية ونور العالمين، سيدنا محمد (صل الله عليه وآله وسلم)

الى من كلله الله بالهيبة والوقار.. الى من احمل اسمه بكل افتخار الى من علمني العطاء دون انتظار ..... والدي العزيز

الى منبع الحنان، الى من كان دعائها سر نجاحي وحنانها بلسم لجرحي الى أغلى الاحباب ..... أمي

الى القلوب البريئة، الى رياحين حياتي .....اخوتي واخواتي

الى من مهد طريق العلم .....معلماتي ومدرساتي واستاذاتي واساتذتي

جدول المحتويات

I	الخلاصة	
1	**الفصل الاول** **التعلم الآلي والتعلم العميق**	1
1	المقدمة	1.1
2	تعلم الآلة	2.1
2	البرمجة التقليدية مقابل التعلم الآلي	3.1
3	مناهج التعلم الآلي	4.1
4	التعلم العميق	5.1
5	تاريخ التعلم العميق	6.1
6	كيف يعمل التعلم العميق	7.1
6	مقارنة بين التعلم الآلي والتعلم العميق	8.1
7	تطبيقات التعلم العميق	9.1
8	**الفصل الثاني** **التعلم العميق الخاضع للإشراف**	2
8	الشبكات العصبية الاصطناعية	1.2
9	بيرسيبترون	2.2
10	خوارزمية التعلم بيرسيبترون	3.2
12	مسألة XOR وعدم قدرة بيرسيبترون على حلها	4.2
13	شبكات التغذية العميقة	5.2
14	دالة التنشيط	6.2
17	الاستنتاجات	7.2
18	المصادر	

الخلاصة

يتمثل الهدف الأساسي للذكاء الاصطناعي في تمكين الآلات من محاكاة السلوك البشري وحل المشكلات بذكاء، وهو ما يتحقق بفعالية من خلال تقنيات التعلم الآلي والتعلم العميق .يُعرف التعلم في هذا السياق بأنه تحسين أداء الآلة في مهمة معينة باستخدام الخبرة والبيانات دون برمجة صريحة، وينقسم منهجياً إلى تعلم خاضع للإشراف، وغير خاضع للإشراف، ومعزز .ويمثل التعلم العميق تطوراً متقدماً في هذا المجال؛ إذ يعتمد على الشبكات العصبية الاصطناعية المستوحاة من البنية البيولوجية للدماغ البشري، ويتفوق بقدرته الفائقة على الاستخراج التلقائي للسمات المعقدة من البيانات الضخمة دون الحاجة إلى هندسة الميزات يدوياً . تتكون هذه الشبكات من طبقات متعددة (إدخال، مخفية، ومخرجات) تعالج البيانات وتتعلم تدريجياً من خلال التعديل الرياضي المستمر للأوزان المرتبطة بالخلايا العصبية .وفي حين يُعد "البيرسيبترون " (Perceptron)أبسط أشكال هذه الشبكات كمصنف ثنائي، إلا أن قصوره الرياضي في حل المشكلات غير القابلة للفصل خطياً، مثل مسألة(XOR) ، أدى إلى الاعتماد الواسع على شبكات التغذية الأمامية العميقة متعددة الطبقات .ولضمان كفاءة هذه المعمارية المعقدة، تلعب دوال التنشيط غير الخطية، مثل (Sigmoid) و(ReLU)، دوراً حاسماً ورياضياً في تحديد حالة تنشيط الخلايا العصبية ونقل الإشارات بين الطبقات، مما يمنح نماذج التعلم العميق مرونة وقوة حوسبية استثنائية لحل أعقد المسائل الحديثة.

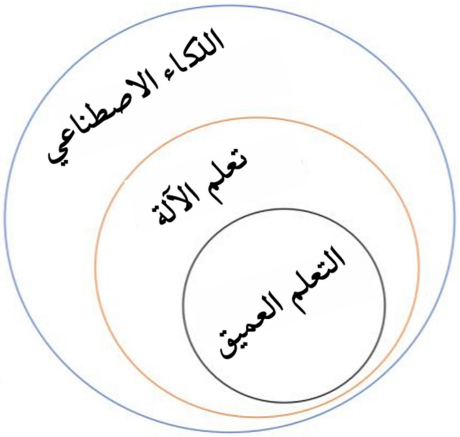
الفصل الأول

التعلم الآلي والتعلم العميق

١.١ المقدمة

الهدف الأساسي في مجال الذكاء الاصطناعي هو إعطاء أجهزة الكمبيوتر القدرة على فهم العالم من حولهم والتفاعل معه بطريقة ذكية. على مدى السنوات العديدة الماضية، برز التعلم العميق كواحد من أكثر الأساليب الواعدة لتحقيق هذا الهدف.(١)

التعلم العميق هو طريقة للتعلم الحسابي للمفاهيم عالية المستوى في البيانات وتمثيلها باستخدام شبكات عصبية هرمية عميقة وهي جزء من أساليب التعلم الآلي.



الشكل ١- م العميق.

الذكاء الاصطناعي: كما يوحي الاسم، [illegible] الذكاء البشري في الآلة، بطريقة تحاكي السلوك الشبيه بالإنسان وتحل المشكلات بشكل إبداعي. بتعبير أدق، يحاول الذكاء الاصطناعي إيجاد طريقة لتنفيذ نسخة من الدماغ البشري، أي كما يفكر الإنسان ويعمل. (٢)

التعلم الآلي: التعلم الآلي، وهو مجموعة فرعية من الذكاء الاصطناعي، يمكن الكمبيوتر ليكون قادرا على التعلم من خلال التجربة دون تخطيط صريح.

التعلم العميق: التعلم العميق هو مجموعة فرعية من التعلم الآلي ويستخدم بنية الشبكات العصبية لتقليد عملية صنع القرار لحل مشكلة مشابهة للدماغ البشري، ويقوم بنفس الشيء مثل التعلم الآلي، ولكن بقدرات مختلفة. بمقارنة التعلم الآلي مع التعلم العميق يمكن القول إنه بينما يستخرج التعلم العميق الميزات تلقائيًا من بنية البيانات، يجب أن يتم ذلك يدويًا عن طريق التعلم الآلي. وإذا قام بتنبؤات خاطئة في تقرير حل المشكلة، فإن الخبير أو المبرمج يجب أن يحل المشكلة بشكل صريح. لذلك يمكن اعتبار التعلم العميق نسخة متطورة ومتقدمة من التعلم الآلي. (٣)

يمكن تعريف التعلم على أنه تحسين الأداء في مهمة معينة باستخدام الخبرة والممارسة. يتم تصنيف السلوك الذكي للإنسان من خلال التعلم في التجارب، والتعلم هو مصدر المرونة في حياة الفرد.

إنّ القدرة على التعميم لتطبيق المعرفة المكتسبة على أمثلة لم يرها من قبل هي السمة الرئيسية للتعلم البشري والآلي بالطبع، يعد التعلم البشري شيئًا أكثر تقدمًا حتى من خوارزميات التعلم الآلي الأكثر تقدماً. (٤)

٢.١ تعلم الآلة

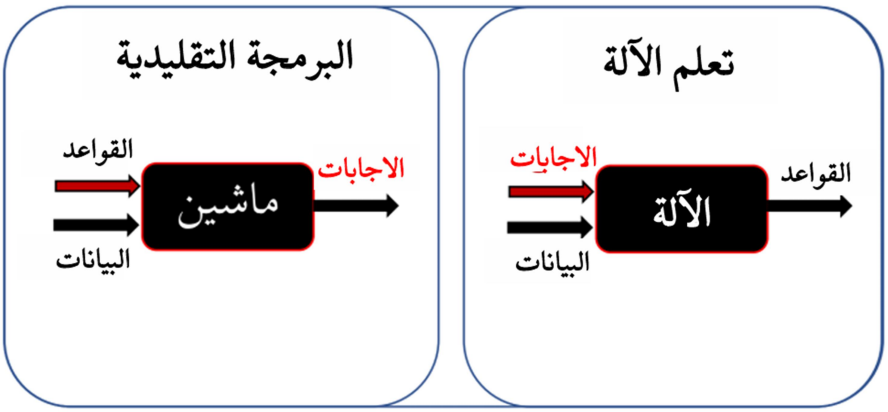
التعلم الآلي هو فرع من فروع الذكاء الاصطناعي حيث تتعلم الآلة أداء المهام التي لم يتم التخطيط لها بشكل صريح، ويتم تحسين أداء الجهاز تلقائيا من خلال الخبرة في أداء هذه المهمة. لقد قطعت محاولة إنشاء آلة تفكر مثل الإنسان شوطاً طويلاً. لقد تحول التعلم الآلي من نظام غامض إلى قوة صناعية واجتماعية رئيسية في صنع القرار الآلي، من الأعمال التجارية عبر الإنترنت والإعلان إلى التعليم والرعاية الصحية. أصبح التعلم الآلي تقنية عامة قوية للعالم نظرًا لقدرته القوية على التعلم من خلال التكيف مع البيانات المصنفة وغير المصنفة. (١)

تبدأ عملية التعلم الآلي باستخدام البيانات الأولية لاستخراج معلومات مفيدة للمساعدة في اتخاذ قرارات أفضل

تستخدم خوارزميات الكمبيوتر بناءً على بيانات التدريب المعطاة لنموذج التعلم تلقائيا من خلال التجربة والاختبار من البيانات، بحيث في مواجهة البيانات الجديدة في نفس المجال، يمكن أن تظهر سلوكاً شبيها بالبشر. يمكن تنفيذ التعلم الآلي وتنفيذه بأساليب مختلفة مناهجها الرئيسية الثلاثة هي التعلم الخاضع للإشراف والتعلم غير الخاضع للإشراف والتعلم المعزز.

٣.١ البرمجة التقليدية مقابل التعلم الآلي

عندما يتعلق الأمر باستخدام أجهزة الكمبيوتر لأداء مهام بشرية، يجب علينا دائما إعطاء التعليمات للكمبيوتر في شكل برنامج كمبيوتر. لغات البرمجة التقليدية هي عملية يدوية (بمعنى أنه يجب على المبرمج إنشاء البرنامج) وعادة ما يأخذ البيانات ومجموعة من القواعد كمدخلات، ومن خلال تطبيق هذه القواعد على البيانات، يحصلون على الإجابات كمخرجات في التعلم الآلي، من ناحية أخرى، يتم تقديم البيانات والإجابات (أو العلامات) كمدخلات، ويتم استخدام القواعد (النماذج) كمخرجات (الشكل١-٢). نمط التعلم الآلي له قيمة لا مثيل لها. لأنه يسمح للآلة بتعلم قواعد جديدة في مساحة معقدة وكبيرة يصعب على البشر فهمها. (١)



٤.١ مناهج التعلم الآلي

١.٤.١ التعلم الخاضع للإشراف

في فصل الرياضيات يشرف المعلم على كيفية حل مشكلة أو حل مشكلة بشكل صحيح أو خاطئ. هذا الموقف مشابه لما تتبعه خوارزمية الخاضعة للإشراف في وضع التعلم، يراقب النظام التعلم لكل تعليمات إرشادية بناءً على عنصر مزدوج من المدخلات والتسمية، وهو ناتج المشكلة تعني مجموعة البيانات الموسومة أن كل عضو في مجموعة التدريب يحصل على إجابة أو حل. الهدف هنا هو تكييف النظام بطريقة تمكن إدخال النظام الجديد من توقع المخرجات الصحيحة بناءً على ما تعلمته حتى الآن من بيانات التدريب في التعلم الخاضع للإشراف، إذا كانت بيانات مسئلة التعلم منفصلة، فهي مسئلة تصنيف، وإذا كانت قيم البيانات مستمرة، فإنها تسمى الانحدار (التوقع). (١)

٢.٤.١ التعلم غير الخاضع للإشراف

يحدث التعلم بدون إشراف بدون مساعدة مشرف او مراقب، تماما كما تتعلم السمكة السباحة بمفردها. في التعلم غير الخاضع للإشراف، تكون خوارزمية التعلم مخصصة المدخلات محددة، بدون مخرجات محددة. الهدف من هذا التعلم هو العثور على هذه المخرجات بمفردها، والتي يتم إجراؤها بواسطة الخوارزمية نفسها عن طريق تحليل البيانات وتحديد الأنماط المخفية والمحددة الموجودة في بنية البيانات التجميع (التكتل) هو مثال على هذا النوع من التعلم، والذي يهدف إلى اكتشاف مجموعات من العناصر المتشابهة بناءً على التشابه المقاس، أو لفهم أوجه التشابه في بنية البيانات. (٢)

٣.٤.١ التعلم المعزز

التعلم المعزز يحل المشكلة بطريقة مختلفة. يعتمد التعلم المعزز على التفاعل مع البيئة وتتعلم الخوارزمية التفاعل مع البيئة وحدها. في هذا النوع من التعلم، تتعلم الخوارزمية من خلال آلية التغذية الراجعة والتجارب السابقة وتحاول حل المشكلات بنفس الطريقة التي يمثل بها البشر في الحياة، وتتعلم كيفية تحسين سلوك الوكيل بناءً على وجود أو عدم وجود المكافآت هدفها هو العثور على مجموعة من أنماط الإجراءات، عن طريق اختبارها ومقارنتها جميعًا للحصول على أعلى درجة مكافأة. لا يتطلب هذا النوع من التعلم مجموعة بيانات تعليمية. بمعنى آخر، إنه ليس تعلما خاضعا للإشراف ولا تعلما غير خاضع للإشراف. (٢)

المؤشر	التعلم الخاضع للإشراف	التعلم غير الخاضع للإشراف	التعلم المعزز
التعريف	يتعلم من خلال مجموعات البيانات ذات العلامات.	يتم تدريبه بدون توجيه من خلال البيانات غير المسماة.	يتفاعل مع بيئة العمل.
انواع البيانات	البيانات المصنفة	بيانات غير المصنفة	لا يوجد تعريف بيانات
نوع المشكلة	التصنيف والانحدار	قواعد الرابطة والتكتل (التجميع)	معتمد على المكافآت
المراقب	مراقب إضافي	بدون مراقب	بدون مراقب
الهدف	تعيين بيانات الإدخال لمخرجات محددة	اكتشاف النمط	تعلم سلسلة من الإجراءات

٥.١ التعلم العميق

يزن دماغ الطفل أقل من نصف كيلوغرام. ومع ذلك، فهو قادر على حل المشكلات التي لا تمتلكها حتى أكبر وأقوى أجهزة الكمبيوتر العملاقة من صنع الإنسان. بعد عدة أشهر من الولادة، يمكن للطفل التعرف على وجوه والديه، وتحديد الأشياء المتعارضة، وحتى تمييز الأصوات في الطفولة المبكرة، كان لديهم فهم واضح للقواعد وحفظوا آلاف الكلمات. (١)

على مدى السنوات القليلة الماضية، كان البشر يحاولون بناء آلات ذكية مثل الروبوتات التي تنظف المنازل، والسيارات ذاتية القيادة، والأنظمة التي تكتشف الأمراض تلقائيًا. يتطلب بناء مثل هذه الآلات الاصطناعية الذكية حل بعض أكثر المشكلات الحسابية تعقيدا التي واجهناها على الإطلاق؛ المشاكل التي يمكن للدماغ البشري حلها في أجزاء من الثانية. (٣)

التعلم العميق هو مجموعة فرعية من التعلم الآلي ويركز على دراسة وتطوير الآلات التي يمكنها التعلم. في تعريف أكثر دقة التعلم العميق مع معالجة البيانات والمشابه للإنسان، من خلال تعلم مثال تم تعلمه مسبقاً، ومحاولة استخراج ميزات محددة تلقائيا، من خلال عدد طبقات التسلسل ما هو موجود في هيكلها هو إنشاء نموذج لاتخاذ القرار من أجل حل مشكلة. يسمح وجود هذا العدد من الطبقات المختلفة بالتعلم العميق ليكون

قادرا على اكتشاف ميزات محددة للمشكلة في كل طبقة واستخدامها لاتخاذ قرارات أفضل في حل المشكلة. (٣)

٦.١ تاريخ التعلم العميق

التعلم العميق موجود منذ الأربعينيات من القرن الماضي تحت عناوين مختلفة مثل: السبرانية والاتصالية والأكثر شهرة، الشبكات العصبية.

تم تقديم النموذج الأول للشبكات العصبية بواسطة ماك كلاج وبيتزا في عام ١٩٤٣. كانت الشبكة عبارة عن مصنف ثنائي يمكنه التمييز بين فئتين مختلفتين بناءً على قيم الإدخال كانت مشكلة هذه الشبكة هي تعديل الأوزان بواسطة عامل بشري. بعد ذلك، في عام ١٩٥٧، اقترح روز نبلات خوارزمية بيرسيبترون (Perceptron)، والتي يمكن أن تتعلم الأوزان لتصنيف البيانات في هيكلها دون تدخل عامل بشري. بينما تم استخدام طريقة البيرسبترون لعدة سنوات، في عام ١٩٦٩ نشر مينسكي وبوبيرت مقالا يوضحأن بيرسيبترون لا يمكنها إلا تصنيف المشكلات الخطية، وأن المشكلات غير الخطية لا يمكن حلها بهذه الطريقة. ادعى مؤلفو هذه المقالة أيضا في نفس العام أنه لا توجد موارد حسابية مطلوبة لبناء شبكات عصبية كبيرة وعميقة، وهو ادعاء أدى إلى تدمير المقالات على الشبكات العصبية. لحسن الحظ، أدى إدخال خوارزميات الانتشار الخلفي بواسطة فيربوس ٣ (١٩٧٤) و روملهات (١٩٨٦) و لكان (١٩٩٨) إلى الإحياء المبكر للشبكات العصبية. في هذا البحث، كانوا قادرين على تدريب شبكة عصبية متعددة الطبقات. (٤)

اليوم، تعد خوارزميات الانتشار الخلفي أساس الشبكات العصبية، والتي يمكننا من خلالها تدريب الشبكة، وكذلك التعلم من أخطائها. لكن في ذلك الوقت، بسبب ضعف أجهزة الكمبيوتر ونقص مجموعات البيانات الكبيرة، لم يتمكنوا من تدريب الشبكات العصبية بأكثر من طبقتين مخفيتين. ولكن اليوم، مع تزايد قوة الأجهزة وعصر البيانات الضخمة، والتي توفر الكثير من البيانات لتدريب الشبكة، يمكن تعليم الشبكات التي تحتوي على أكثر من بضع طبقات مخفية. تسمى الشبكات العصبية المكونة من عدة طبقات الشبكات العميقة. عندما نستخدم الشبكات العميقة اليوم، فإننا نعني التعلم العميق. (٤)

٧.١ كيف يعمل التعلم العميق؟

تتعلم نماذج التعلم العميق من خلال التحليل المستمر للبيانات واكتشاف الهياكل المعقدة في البيانات. تتحقق عملية التعلم من خلال بناء نماذج حسابية تسمى الشبكات العصبية المستوحاة من بنية الدماغ. يتكون هيكل هذه الشبكة من عدة طبقات معالجة. (١)

يسعى التعلم العميق إلى استغلال البنية غير المعروفة في توزيع المدخلات من أجل اكتشاف تمثيل جيد من خلال هيكل هرمي للمفاهيم التي تشبه طبقات المعالجة.

في هذا الهيكل، بالانتقال إلى طبقات المستوى التالي، يكون قادرا على حل المفاهيم الأكثر تعقيداً للمشكلة. تقوم الطبقات الأولية بمعالجة البيانات الأولية والطبقات اللاحقة قادرة على استخدام المعلومات العصبية الموجودة في الطبقات السابقة للحصول على معلومات أكثر تعقيداً من البيانات.

في التعلم العميق يتم استخراج الميزات داخل النموذج دون تدخل بشري لعمل التنبؤ المطلوب للمشكلة.

الميزة الثانية للتعلم العميق هي أنه يتم تزويده بكمية كبيرة من البيانات. تميل نماذج التعلم العميق إلى زيادة الدقة مع زيادة كمية البيانات التعليمية. في المقابل، لا تتحسن نماذج التعلم الآلي التقليدية بعد نقطة تشبع واحدة. (٣)

الشكل ١-٣ الفرق بين التعلم الآلي والتعلم العميق في استخراج الميزات

٨.١ مقارنة بين التعلم الآلي والتعلم العميق

من أهم الاختلافات التي تقارن التعلم العميق بالتعلم الآلي هو أداء النظام بناءً على زيادة عدد الأمثلة التدريبية لن يحقق التعلم العميق نتائج جيدة في حالة عدم وجود أمثلة تدريب كافية. في المقابل، يمكن أن يُظهر التعلم الآلي نتائج جيدة حتى مع وجود عدد صغير من الأمثلة. يتطلب استخدام التعلم العميق أيضًا أجهزة متقدمة، بينما يمكن استخدام التعلم الآلي مع كل من الأجهزة وأجهزة الكمبيوتر منخفضة الطاقة.

يتمثل الاختلاف المهم والرئيسي الذي يظهر قوة التعلم العميق مقابل التعلم الآلي في الاستخراج التلقائي للميزات في هذه الخوارزميات باختصار يمكن رؤية المقارنة بين التعلم الآلي والتعلم العميق في الجدول ١-٢. (٢)

المعايير	التعلم الآلي	يادگيرى عميق
هندسة الميزات	تحتاج إلى فهم الميزاتفي البيانات	لا حاجة لاستخراج الميزات يدويًا
الاعتماد على البيانات	أداء ممتاز على البيانات المنخفضة والمتوسطة	أداء رائعفي البيانات الضخمة
الاعتماد على الأجهزة	يعمل على الاجهزة الضعيفة أيضًا	يحتاج الى اجهزة قوية
وقت التنفيذ	من بضع دقائق إلى بضع ساعات	في بعض الأحيان تصل إلى بضعة أسابيع

٩.١

- [illegible]
- الروبوتات
- المساعدين الافتراضيين
- الأتمتة الصناعية
- الزراعة
- تحديد الأنواع البحرية
- البحث الطبي

- التصوير الطبي
- أنظمة التوصية
- التعرف على الإيماءات والحركة
- تشخيص تأخر النمو عند الأطفال

الفصل الثاني

التعلم العميق الخاضع للإشراف

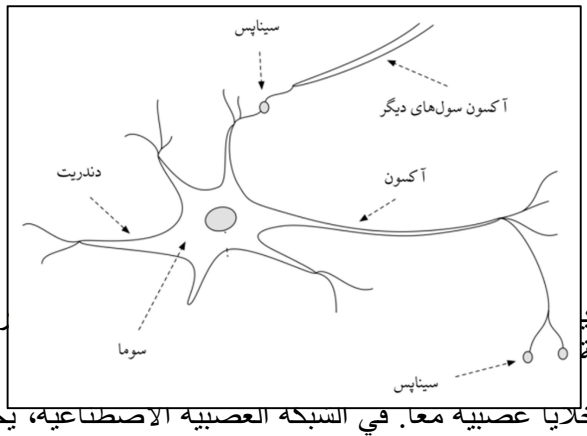
١.٢ الشبكات العصبية الاصطناعية

على الرغم من أن الدماغ هو أحد أكثر المكونات تعقيداً وقوة في بنية الكائنات الحية، إلا أن بنيته الأساسية بسيطة نسبيا ومفهومة الشبكات العصبية الاصطناعية هي نسخ من هذا الهيكل الشبكات العصبية الاصطناعية هي نماذج حسابية تحاكي آلية التعلم لبنية الدماغ البشري مثل الشبكة العصبية الطبيعية.

تتكون بنية الدماغ البشري، وهي الشبكة العصبية الطبيعية، من عدد كبير من الوحدات البسيطة تسمى الخلايا العصبية. تحتوي الخلايا العصبية على ثلاثة أنواع من المكونات التشعبات، والسوما، والمحاور. يمكن رؤية منظر لها في الشكل ٢-١. (٣)

تستخدم الخلايا العصبية المحاور والتشعبات للتواصل مع بعضها البعض. تسمى مناطق الربط بين المحاور والتشعبات نقاط الاشتباك العصبي غالبا ما تتغير هذه النقاط استجابة للمحفزات الخارجية، وهذه التغييرات هي التي تؤدي إلى التعلم في الكائنات الحية. يولد كل خلية عصبية إشارة داخل نفسها عن طريق إنشاء عمليات كيميائية، ويتم إرسال الإشارات المرسلة من كل خلية عصبية بواسطة المحور العصبي إلى التشعبات العصبية الأخرى بناءً على هذه الإشارات، يقوم الدماغ بعمل محدد. الشبكات العصبية الاصطناعية هي محاكاة لبنية الدماغ البشري وتستند إلى هذه الفرضيات: (٢)

١. تتم معالجة المعلومات في هياكل بسيطة بأعداد كبيرة تسمى الخلايا العصبية.
٢. تنتقل الإشارات من خلال الاتصالات بين الخلايا العصبية في الشبكة.
٣. كل اتصال له وزنه الخاص، والذي يتم ضرب هذه الأوزان في الشبكة العصبية بواسطة إشارة الإرسال.
٤. يستخدم كل خلية عصبية دالة تنشيط لتطبيقها على مدخلاتها، وهي مجموع الاوزان الإشارات الإدخال، لتوليد إشارة الخرج.



الشكل ٢-١ منظر لشبكة عصبية طبيعية

وفقا لهذه الفرضيات، فإن نمط الاتصال بين الخلاي ريقة تحديد الأوزان على الاتصالات تسمى الخوارزمية

تشكل الشبكة العصبية طبقة من خلال ضم عدة خلايا عصبية معا. في الشبكة العصبية الاصطناعية، يحسب حساب الدالة عن طريق نشر القيم المحسوبة من الخلايا العصبية المدخلة إلى الخلايا العصبية الناتجة واستخدام الأوزان كمعاملات وسيطة. يحدث التعلم عن طريق تغيير الأوزان المرتبطة بالخلايا العصبية. وفقا للمحفزات الخارجية اللازمة لتعلم بنية الدماغ، توجد بيانات تعليمية في الشبكات العصبية الاصطناعية لهذا المحفز الخارجي، وهي أمثلة على أزواج المدخلات المخرجات. (١)

من خلال ضبط الأوزان بشكل تسلسلي بين الخلايا العصبية بناء على مدخلات ومخرجات بيانات التدريب، يتم تحسين الأداء المحسوب بواسطة الشبكات العصبية بمرور الوقت لعمل تنبؤات أكثر دقة. لذلك، إذا تم تدريب الشبكة العصبية بالعديد من الصور المختلفة، فستتمكن في النهاية من التعرف على صورة القط بشكل صحيح في الصورة التي لم ترها من قبل. (٤)

يتشكل سلوك الشبكة العصبية من خلال بنية تلك الشبكة. يتم تعريف هذه البنية على أساس ما يلي:

- عدد الخلايا العصبية.
- عدد الطبقات.
- كيفية الاتصال بين الطبقات.

يُطلق على البنية الأكثر شهرة للشبكة العصبية اسم الشبكة العصبية متعددة الطبقات، والتي تسمى أيضا بيرسيبترون متعدد الطبقات، والذي يتكون من ثلاث طبقات طبقة الإدخال، والطبقة المخفية، وطبقة الإخراج. تتلقى طبقة الإدخال المعلومات، تقوم طبقة مخفية واحدة أو أكثر بتنفيذ عملية المعالجة، وتعرض الطبقة الناتجة من النتائج. مع زيادة عدد الطبقات المخفية، ننتقل إلى شبكة أعمق لديها القدرة على حل مشاكل أكثر تعقيداً من نظيراتها الضحلة. فيما يلي، سوف نقدم الشبكات العصبية أحادية الطبقة ومتعددة الطبقات. (٢)

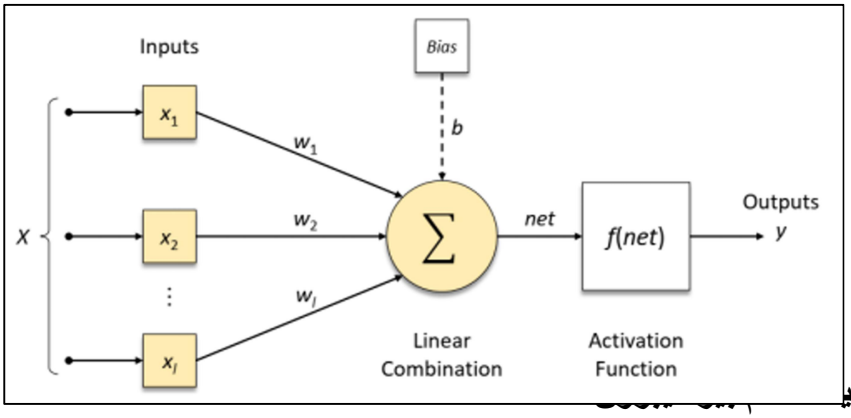
٢.٢ بيرسيبترون

اخترعت بيرسيبرون بواسطة فرانك روزنبلات في عام ١٩٥٧ في مختبر كورنيل للطيران، بيرسيبرون هو أبسط شكل للشبكة العصبية الاصطناعية وهو عبارة عن مصنف ثنائي. إن بُنية هذه الشبكة العصبية ليست سوى طبقة إدخال واحدة بمخرج واحد فقط، ومن ثم يطلق عليها أيضًا اسم الشبكة العصبية أحادية الطبقة. يظهر مثال على ذلك في الشكل ٢–٢. كما يمكن رؤيته، هناك عدد كبير من المدخلات في هذه الشبكة، والتي يتنبأ مجموعها، بعد الحساب، بالإخراج باستخدام دالة التنشيط. مع قائمة المدخلات $x = \{x_1, x_2, \dots, x_n\}$. سيكون لكل إدخال متجه وزن $w = \{w_1, w_2, \dots, w_n\}$. يتم حساب مجموع الاوزان بالمعادلة(1). (١)

$$net = \sum_{i=0}^{n} {}_i \cdot x_i + b \qquad (1)$$

ثم تحصل دالة التنشيط على الناتج y بناءً على العتبة :θ

$$y = f(net) = \begin{cases} 1, if\ net\ \geq \theta \\ 0, if\ net\ < \theta \end{cases} \qquad (2)$$



3.2 خوارزمي

مع الافتراضات المناسبة، يمكن إثبات أن التعلم في بيرسيبرون سوف يتقارب مع الأوزان الصحيحة من خلال تكرار الخوارزمية الخاصة به. أي أن التعلم الشبكي سيؤدي إلى تقدير الأوزان التي تمكن الشبكة من إنشاء القيم الصحيحة في المخرجات بناءً على مدخلات المشكلة. في خوارزمية بيرسيبرون، لكل متجه إدخال

أثناء التدريب، فإنه يولد شبكة الإخراج ويقارنها بالقيمة الصحيحة لتحديد ما إذا كان قد حدث خطأ. في هذه الشبكة، إذا لم يحدث خطأ، فلن تتغير الأوزان وسيستمر التدريب حتى تصبح الشبكة خالية من الأخطاء. (٢)

لنفترض أننا قمنا بتغذية متجه إدخال إلى الشبكة وأن إحدى الخلايا العصبية تقدم إجابة خاطئة؛ أي أن مقدار المدخلات ليس هو نفسه الهدف. هناك وزن m لهذه الشبكة، لكل خلية عصبية متصلة بعقدة الإدخال تحدد وزناً. إذا أطلقنا على k الخلية العصبية التي أنتجت ناتجاً خاطئاً، فإن الأوزان هي w_{ik}، حيث يمكنني أن يكون من ١ إلى m. بناءً على هذا نعرف الأوزان التي تتغير. الآن دعونا نلقي نظرة على كيفية تغير الأوزان. السؤال الأول هو ما إذا كان كل وزن كبير جداً أم صغيراً جداً. (٢)

حساب $y_k - t_k$ هو الفرق بين الإخراج (y_k) الذي تقوم به الخلية العصبية والقيمة المستهدفة (t_k) التي يجب أن تقوم بها الخلية العصبية. إذا أصبحت هذه القيمة سالبة، فإننا نجعلها أكبر، وعلى العكس، إذا كانت موجبة، فيمكن التحكم فيها عن طريق تقليل الخطأ. تعتمد كيفية تغيير الأوزان على المعادلة التالية:

$$\Delta w_{ik} = -(y_k - t_k) \times x_i$$

قبل أن (3) ينتهي قانون التعلم، ومع ذلك، يجب أن نقرر مدى تغير الأوزان. يتم ذلك بضرب القيمة العالية بواسطة معامل يسمى **معدل التعلم**، والتي يتم تمثيله عادةً بالرمز η. تحدد قيمة معدل التعلم سرعة التعلم للشبكة وهي مهمة للغاية. بناءً على ذلك، نكتب المعادلة الأخيرة لتحسين الوزن على النحو التالي:

$$w_{ij} \leftarrow w_{ij} - \eta(y_i - t_i).x_i$$

بناءً على (4) ذلك، يمكن رؤية خوارزمية تعلم بيرسيبرون على النحو التالي:

خوارزمية التعلم بيرسيبرون

- **تهيئة القيم الاولية**

قم بتعيين قيم صغيرة لجميع الأوزان w_{ij}.

- **التدريب**

لتكرار T حتى تكون جميع النواتج صحيحة:

– حساب دالة التنشيط لكل خلية عصبية j باستخدام وظيفة التنشيط g:

$$y_j = g\left(\sum_{i=0}^{m} w_{ij}x_i\right) = \begin{cases} 1, IF \sum_{i=0}^{m} w_{ij}x_i > 0 \\ 0, IF \sum_{i=0}^{m} w_{ij}x_i \leq 0 \end{cases} \quad (5)$$

– يتم ضبط كل من الأوزا

$$w_{ij} \leftarrow w_{ij} - \eta(y_i - t_i).x_i \quad (6)$$

- **الاستدعاء**

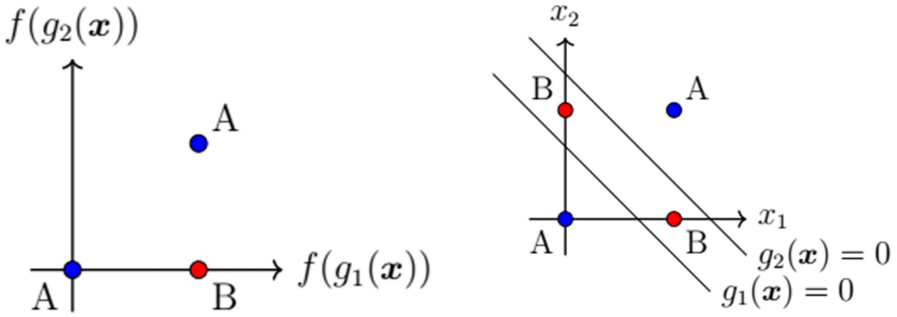
– حساب دالة التنشيط لكل خلية عصبية j:

$$y_j = g\left(\sum_{i=0}^{n} w_{ij}x_i\right) = \begin{cases} 1, \text{اگر } w_{ij}x_i > 0 \\ 0, \text{اگر } w_{ij}x_i \leq 0 \end{cases} \quad (7)$$

4.2 مسألة XOR وعدم قدرة بيرسيبترون على حلها

تقوم خوارزمية التعلم بيرسيبترون بضبط الأوزان بحيث يتم تصنيف جميع عينات الإدخال بشكل صحيح. لن تكون خوارزمية التعلم هذه محدودة إذا لم تكن المدخلات قابلة للفصل خطيًا عن بعضها البعض. لن تكون هذه الشبكة العصبية أحادية الطبقة قابلة للاستخدام في مشاكل العالم الحقيقي **لأن الأنماط بين الفئات ليست بالضرورة قابلة للفصل خطيًا**. لذلك نذهب إلى مجموعة من الشبكات التي لديها قوة أكبر في حل مشاكل العالم الحقيقي. (٣)

يعد استخدام الطبقة المخفية في بنية الشبكات بمثابة هروب من القيود الموجودة في الشبكات أحادية الطبقة. لفهم هذه المسألة بشكل أفضل، ضع في اعتبارك الشكل (أ) ٢–٣؛ حيث [٠،٠] و [١،١] تنتمي إلى الفئة A و [٠،١] و [١،٠] تنتمي إلى الفئة B. من السهل ملاحظة أنه لا توجد خطوط مستقيمة يمكنها فصل الفئتين تمامًا. لذلك، فإن المصنف الخطي مثل بيرسيبترون لديه أداء ضعيف للغاية في هذه الحالة. (٣)

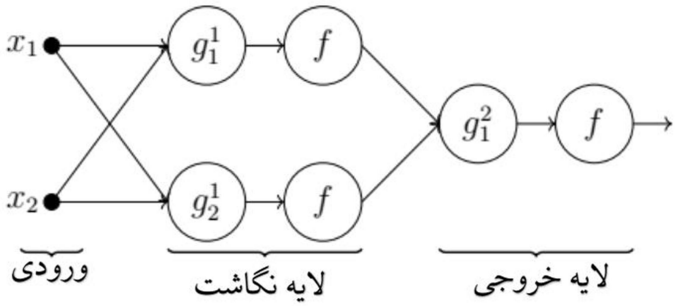


(آ) مسئلة تصنيف XOR (ب) مسئلة XOR في تعيينها

ومع ذلك، م م بسهولة أن المسافة بين g_2 و g_1 مخصصه للفئه B ويجب تخصيص المساحه ادناه g_2 او اعلى g_1 للفئة A. الآن ضع في اعتبارك هذا التعيين:

$$x \rightarrow \begin{bmatrix} f(g_1(x)) \\ f(g_2(x)) \end{bmatrix} \quad , \quad f(x) = \begin{cases} 1, if\ net \geq \theta \\ 0, if\ net < \theta \end{cases} \tag{8}$$

الآن $g_i(x)$ كما هو ب) ٢–٣. كما نرى بوضوح، يمكن فصلا ة على البيرسيبرون ثنائي الأبعاد. سيكون الناتج ٠ أو ١، اعتمادًا على الفئة x التي تنتمي إليها. (٤)

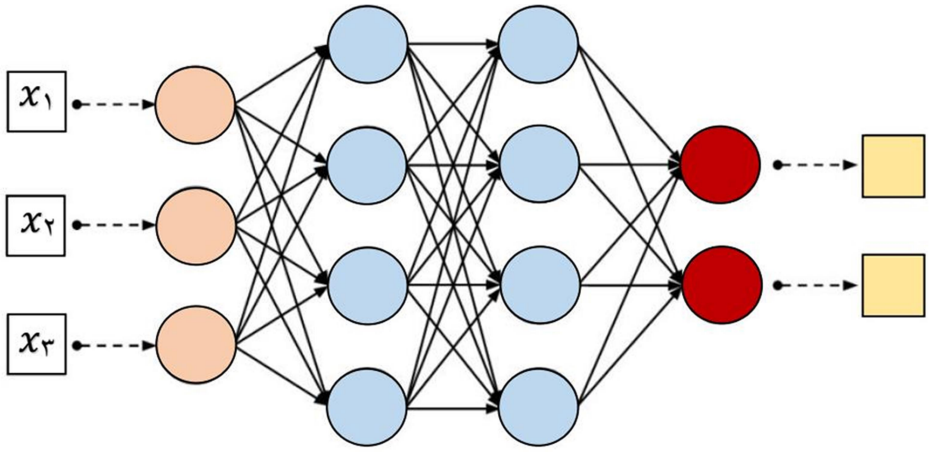


5.2 شبکات التعدیه العمیقه

كان أحد القيود الرئيسية للشبكات أحادية الطبقة أنه لا يمكن تصنيفها إلا بحيث يمكن فصل البيانات خطيًا، وإلا فلن يتم حل مشكلة هذه الشبكات. لحل هذا القيد، يمكن استخدام طبقة مخفية بين طبقات الإدخال والإخراج. ومن الأمثلة على هذه الشبكات، والتي تعد أيضا أساس التعلم العميق الشبكات العصبية بيرسيبترون متعددة الطبقات، والتي يشار إليها أيضا باسم شبكات التغذية العميقة. (٢)

تعد هذه الشبكات واحدة من أكثر الشبكات استخداما في التعلم العميق نظرا لتوافقها مع مجموعة متنوعة من المشكلات. لأنه لا يوجد حد لإدخاله سواء كانت البيانات صورة أو نص أو فيديو سبب تسمية التغذية الامامية هو عدم وجود اتصال تغذية مرتدة، يمكن من خلاله إرجاع مخرجات النموذج إلى النموذج نفسه. تستخدم الخلايا العصبية في كل طبقة دالة تنشيط مشتركة، وبالنسبة لطبقة الإدخال، يكون الإدخال هو المتجه الخام للبيانات. (٢)

في هذه الشبكات، بالانتقال إلى أي طبقة أخرى، يتم حساب مجموع الاوزان للخلايا العصبية للطبقة السابقة وبعد تطبيق دالة التنشيط غير الخطي، يتم نقلها إلى طبقة أخرى للوصول أخيرًا إلى طبقة الإخراج. يوضح الشكل ٢-٥ مخططاً لشبكة عصبية امامية التغذية عميقة متصلة بالكامل.



بناءً على طبقة المخرجات الطبقة المخفية 2 الطبقة المخفية 1 طبقة المدخلات المقادير المدخلة

- طبقة إدخال واحدة فقط.
- طبقة مخفية واحدة أو أكثر، متصلة بالكامل.
- طبقة إخراج واحدة فقط.

ننتقل الآن إلى كيفية عمل كل طبقة من هذه الطبقات وتطبيقها:

طبقة الإدخال هذه الطبقة هي أول طبقة مرئية وتحدد كيفية تلقي بيانات الإدخال (المتجهات) إلى الشبكة. في هذه الطبقة، يمثل عدد الخلايا العصبية عادةً عدد الميزات في الشبكة ترتبط هذه الطبقة تماما بالطبقات المخفية في الشبكات العصبية امامية التغذية. في معماريات الشبكات الأخرى من الممكن أن يكون هذا الاتصال غير متصل بشكل كامل. (١)

الطبقة المخفية: هناك طبقة مخفية واحدة أو أكثر في شبكات العصبية امامية التغذية. قيم الأوزان في اتصالات الطبقة البينية هي كيفية تشفير الشبكات العصبية لاستخراج المعلومات المستفادة من البيانات الأولية. تشير كلمة "مخفية" إلى أن هذه الطبقة غير مرئية للنظام الخارجي. تسمح هذه الطبقة بنمذجة الدوال غير الخطية. سيكون وجود طبقة مخفية كافيًا لمعظم المشكلات؛ كلما زاد عدد الطبقات زاد الوقت الذي يستغرقه إنتاج هذه الشبكة، وبدلاً من ذلك يمكن أن تحل مشاكل أكثر تعقيدا. سيعتمد العدد الأمثل للطبقات والخلايا العصبية في كل طبقة إلى حد كبير على المشكلة. يعد اختيار العدد الصحيح من الخلايا العصبية في كل طبقة مخفية أمرًا بالغ الأهمية، لأنها تلعب دورًا رئيسيا في نجاح عملية حل المشكلات. يجب تحديد هذه الأرقام بشكل تجريبي. سيؤدي اختيار عدد صغير من الخلايا العصبية إلى الضبط الناقص من ناحية أخرى، سيؤدي العدد الكبير من الخلايا العصبية المستخدمة في الطبقة المخفية إلى زيادة غير ضرورية في وقت التدريب وإمكانية الضبط الزائد. (١)

طبقة الإخراج: هذه الطبقة هي آخر طبقة كانت مرئية ويمكن توقع أو حل المشكلة في هذه الطبقة. يمكن أن يعتمد هذا الإخراج على التصميم الذي لدينا للمشكلة، يمكن أن يكون هذا الناتج مجموعة من الاحتمالات

المتعلقة بمشكلة التصنيف، أو ناتج مقيم متعلق بمشكلة الانحدار. يتم تحديد هذا الإخراج بواسطة دالة التنشيط. يتناسب عدد الخلايا العصبية في هذه الطبقة مع المشكلة. على سبيل المثال، إذا كانت المشكلة هي التصنيف الثنائي، فستحتوي طبقة الإخراج على خليتين عصبيتين. (١)

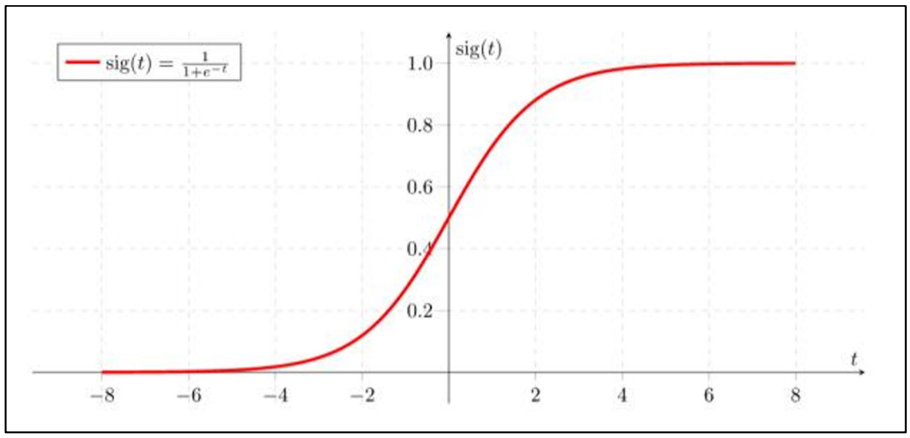
٦.٢ دالة التنشيط

تلعب دالة التنشيط دورًا مهما ورئيسيًا في بنية نموذج الشبكة العصبية. تستخدم هذه الدالة لنشر ناتج كل طبقة إلى أخرى في نهاية العملية الحسابية في كل خلية عصبية. ببساطة، فإن دالة التنشيط هي المسؤولة عن تحديد الخلايا العصبية التي يجب تنشيطها أو التي يجب تعطيلها بشكل عام، يتم استخدام دوال التنشيط غير الخطية بشكل أكثر شيوعًا في الشبكات العصبية. (٢)

١.٦.٢ دالة سيگمويد (sigmoid): تعد دالة سيگمويد، والمعروفة أيضًا باسم الدالة المنطقية، واحدة من أكثر دوال التنشيط غير الخطية فائدة في الشبكات العصبية الاصطناعية. تُستخدم هذه الدالة لحساب احتمال مشاكل التصنيف الثنائي في طبقة المخرجات. تولد هذه الدالة قيم خرج احتمالية بين صفر وواحد لكل فئة، ويمكن تحديدها بواسطة المعادلة (9):

$$\sigma(x) = sigmoid(x) = \frac{1}{1 + e^{-x}} \tag{9}$$

يوضح الشكل ٢–٦ مخططًا لهذه الدالة.



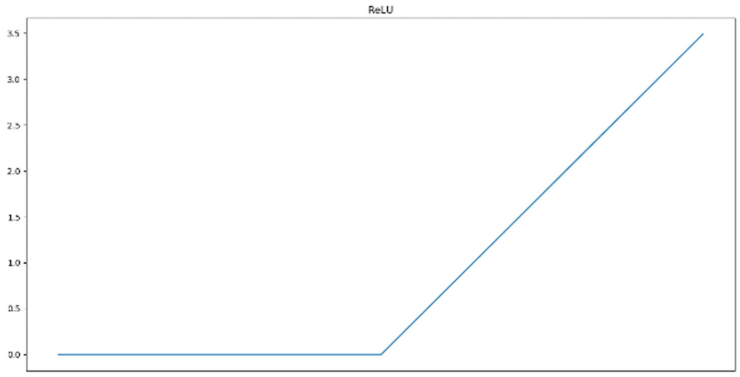
٢.٦.٢ دالة تانتش الزائدية (Tanh): ميزة هذه الدالة هي أنها يمكن أن تتعامل مع الأرقام السالبة بسهولة أكبر. ناتج هذه الدالة هو قيمة بين ١ و _1 ، ويمكن عرضه في صورة المعادلة (10):

$$\sigma(x) = tanh(x) = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}} \tag{10}$$

3.6.2دالة الوحدة الخطية المصححة($ReLU$) تعد دالة الوحدة الخطية المصححة، المستخدمة في الطبقة المخفية، واحدة من أكثر الدوال المستخدمة على نطاق واسع في التعلم العميق اليوم. يمكن تمثيل هذه الدالة في المعادلة (١١):

$$ReLU(x) = \max(0, x) \tag{11}$$

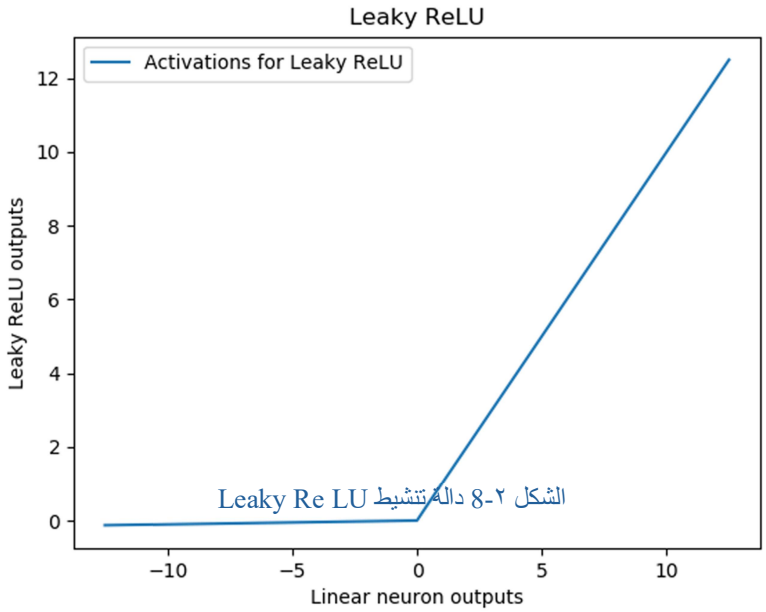
يضبط قيم الإدخال السالبة إلى الصفر لتحويل الإدخال إلى قيمة أكبر من أو تساوي الصفر. بمعنى آخر، لها قيمة صفرية للمدخلات السلبية وقيمة عالية للمدخلات الإيجابية. يمكن رؤية نظرة عامة على هذه الدالة في الشكل ٢–٧.



٤.٦.٢ دالة (LU تمثل الاختلاف في إدخال معامل α تسمح بتنشيط التدرجات الصغيرة إذا لم يتم تنشيطها. لذلك، فإن دالة التنشيط هذه تقضي على مشكلة موت الخلايا العصبية أثناء عملية التدريب. يمكن تمثيل هذه الدالة في المعادلة (12):

$$\text{LeakyReLU}(x) = \begin{cases} x & \text{اگر : صفر} \geq x \\ \alpha x & \text{اگر : صفر} > x \end{cases}$$

يمكن رؤية (2) نظرة عامة على هذه الدالة في الشكل ٢–٨.



الشكل ٢-8 دالة تنشيط Leaky Re LU

٧.٢ الاستنتاج

وفي ختام البحث يمكن ان نستنتج ما يأتي:

- يمكن اعتبار التعلم بمثابة تحسين للأداء في مهمة معينة باستخدام الخبرة والممارسة.
- أي برنامج كمبيوتر يعمل على تحسين أدائه في وظيفة معينة ذات خبرة قد تعلم.
- يمكن تنفيذ التعلم الآلي من خلال ثلاثة مناهج مختلفة التعلم الخاضع للإشراف والتعلم غير الخاضع للإشراف والتعلم المعزز.
- يركز التعلم الآلي على تصميم النماذج التي، في مجال معين، تحصل تلقائيا على خوارزميات الكمبيوتر بناءً على بيانات التدريب المعطاة لنموذج التعلم من خلال التجربة وتجريب البيانات لمواجهة بيانات جديدة في نفس المجال يمكن أن يتصرف المجال بشكل مشابه للبشر.
- يمكن أن يقوم التعلم العميق باستخراج الميزات تلقائيا بدلاً من التعلم الآلي.

- يحدث التعلم في الشبكة العصبية عن طريق تغيير الأوزان المرتبطة بالخلايا العصبية.
- أبسط شكل هو الشبكة العصبية بيرسيبترون وهو مصنف ثنائي.
- تعد الشبكات امامية التغذية العميقة واحدة من أكثر شبكات التعلم العميق استخداما نظرا لعدم وجود قيود على المدخلات.
- تلعب دالة المنشط دورًا مهما ورئيسيًا في بنية الشبكة العصبية.
- تحدد دالة التنشيط الخلايا العصبية التي يجب أن تكون نشطة أو غير نشطة.

المصادر:

١. ميلاد وزن، التعلم العميق المبادئ والمفاهيم والاساليب، ترجمة: الدكتور علاء طعيمه

٢. اوستن زانخ وارخرون، التعمق في التعلم العميق، ترجمة: الدكتور علاء طعيمه،

3. Eli David, Deep Learning, John Wiley & Sons, Inc., New Jersey, 2018.

4. Nithin Buduma, Nikhil Buduma and Joe Papa, Fundamentals of Deep Learning, Second Edition, Nithin Buduma and Mobile Insights Technology Group, LLC, 2022.