وزارة التعليم العالي والبحث العلمي
جامعة أم البواقي
كلية العلوم الاقتصادية والعلوم التجارية وعلوم التسيير
قسم: العلوم الاقتصادية **مقياس:** تطبيقات التعلم العميق

الدرس 04: الشبكات العصبية الأمامية

الشبكات العصبية الأمامية:

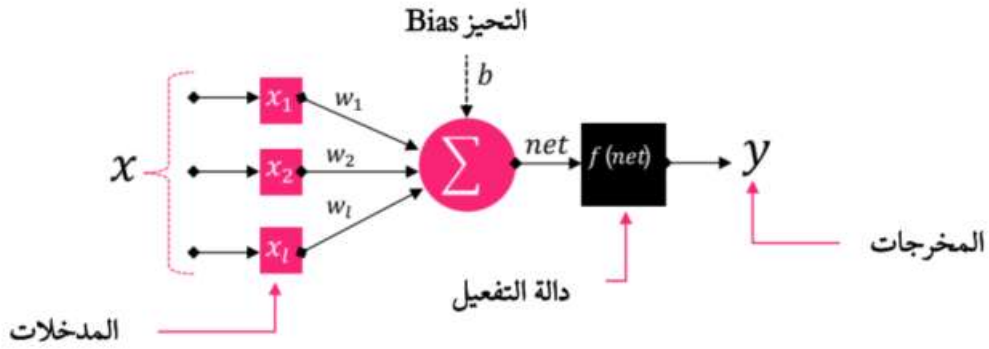
الشبكات العصبية الاصطناعية، أو باختصار، ANNهي مجموعة من نماذج التعلم الآلي المستوحاة بشكل عام من دراسات الجهاز العصبي المركزي للثدييات. بمعنى آخر، إنها نموذج حسابي يحاكي كيفية عمل الخلايا العصبية في دماغ الإنسان. تتكون كل شبكة عصبية اصطناعية من عدة "عصبونات" متصلة منظمة في "طبقات". ترسل الخلايا العصبية في كل طبقة رسائل إلى الخلايا العصبية في الطبقة التالية

تحتوي الشبكة العصبية الاصطناعية على طبقة إدخال وطبقة إخراج وطبقة مخفية واحدة أو أكثر مترابطة. تتكون الطبقة الأولى من الخلايا العصبية المدخلة. ترسل هذه الخلايا العصبية البيانات إلى طبقات أعمق. تتلقى كل طبقة بعد طبقة الإدخال، بدلاً من الإدخال الأولي، إخراج الطبقة السابقة كمدخل. أخيرًا، ينتج آخر طبقة إخراج للنموذج. تحدد أمثلة التدريب بشكل مباشر المخرجات التي يجب إنشاؤها لكل إدخال x. تحاول طبقة المخرجات حساب قيمة قريبة من ناتج معين لعينات تدريب مماثلة. ومع ذلك، فإن سلوك الطبقات الداخلية لا يتأثر بشكل مباشر بعينات التدريب، وخوارزمية التدريب هي التي تحدد كيفية عمل هذه الطبقات من خلال اتخاذ قراراتهم الخاصة لإنتاج المخرجات المرغوبة. ونتيجة لذلك، فإن وظيفة الطبقات الداخلية، بناءً على المخرجات المرغوبة التي تم الحصول عليها من عينات التدريب، غير محددة بوضوح وتعمل كصندوق أسود، ومن ثم تسمى هذه الطبقات بالطبقات المخفية.

لكي تتمكن الشبكات العصبية الاصطناعية من التعلم، يجب أن يكون لديها قدر هائل من البيانات، وهو ما يسمى مجموعة التدريب. عندما تريد تعليم الشبكات العصبية الاصطناعية كيفية التمييز بين القط والكلب، يوفر البرنامج التعليمي آلاف الصور التي تم وضع علامات عليها للكلاب لكي تبدأ الشبكة في التعلم. عند تدريبه على قدر كبير من البيانات، يحاول تصنيف البيانات المستقبلية إلى فئات مختلفة بناءً على ما يعتقد أنه يراه (أو يسمعه، اعتمادًا على مجموعة البيانات). خلال فترة التدريب، تتم مقارنة إخراج الجهاز مع الأوصاف البشرية (العلامات) لما يجب مشاهدته. إذا كانت متطابقة، فإن النموذج يعمل بشكل جيد. إذا كان غير صحيح، فإنه يستخدم الانتشار الخلفي backpropagation لضبط التعلم.

1. بيرسيبترون Perceptron

الخلايا العصبية هي عنصر أساسي في أي شبكة عصبية اصطناعية. يُطلق على أبسط نوع من نمذجة الخلايا العصبية اسم بيرسيبترون ، Perceptronوالذي يمكن أن يحتوي على عدد كبير من المدخلات بمخرج واحد. يوضح الشكل 1 مخططًا لبيرسيبترون. يستخدم Perceptron التعلم الخاضع للاشراف لتصنيف أو توقع المخرجات. يصنف المستشعر أحادي الطبقة البيانات عن طريق رسم حدود القرار decision( ) boundaryباستخدام خط محدد



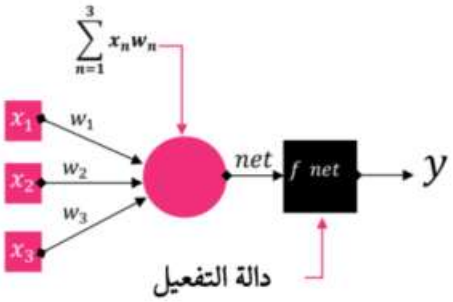
الشكل 1 بيرسيبترون.

دعونا نلقي نظرة على كيفية عمل بيرسبترون. يعمل Perceptronعن طريق التقاط بعض المدخلات العددية جنبًا إلى جنب مع ما يعرف باسم الأوزان weights( ) والتحيزات او الانحيازات bias(.) ثم يقوم بضرب هذه المدخلات بالأوزان المقابلة (المعروفة باسم مجموع الأوزان). ثم يتم إضافة هذا الناتج مع التحيز. تأخذ دالة التنشيط Activation Function() مجموع الاوزان والانحياز كمدخلات وترجع الناتج النهائي. كان الأمر محيرًا !!!! ... دعونا نحلل ، Perceptronلتحسين طريقة عمله. يتكون بيرسيبترون (الشكل 3-1)من أربعة أجزاء رئيسية: قيم الإدخال والأوزان والتحيز (الانحياز) وإجمالي الوزن ودالة التنشيط. افترض أن لدينا خلية عصبية وثلاثة مدخلات $x3$، $x2$، $x1$مضروبة في الأوزان $w3$، $w2$، $w1$على التوالي.

الفكرة بسيطة، بالنظر إلى القيمة العددية للمدخلات والأوزان، فهناك دالة داخل الخلية العصبية تنتج مخرجات. الآن السؤال ما هي هذه الدالة؟ هذه الدالة تعمل كالاتي:

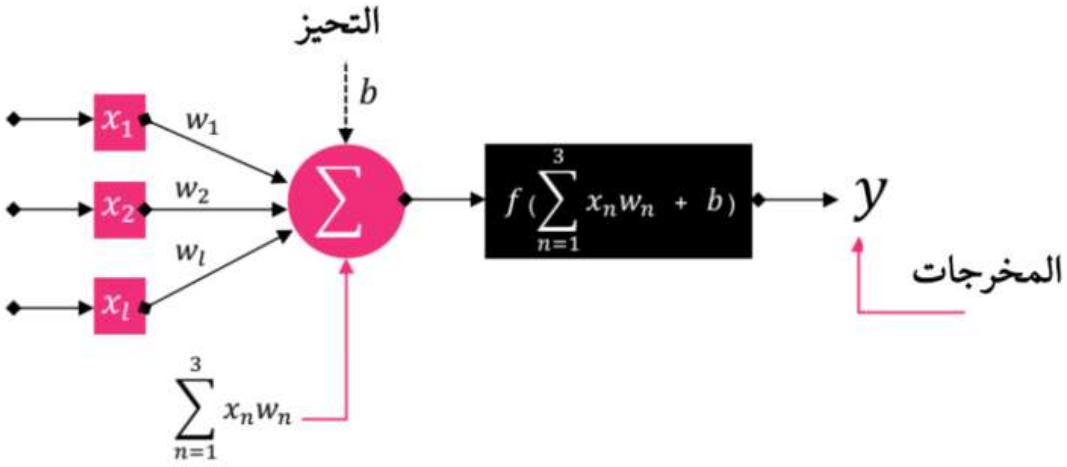
$$y = x_1 w_1 + x_2 w_2 + x_3 w_3$$

هذه الدالة تسمى مجموع الاوزان، لأنها مجموع الأوزان والمدخلات. هذا جيد حتى الآن ، لكن إذا أردنا أن تكون المخرجات في نطاق معين، على سبيل المثال من 0إلى ، 1فماذا يجب أن نفعل؟! يمكننا القيام بذلك باستخدام ما يسمى بدالة التنشيط (دالة التفعيل). دالة التنشيط هي دالة تحول إدخالًا معينًا (في هذه الحالة ، مجموع الاوزان) إلى إخراج محدد ، بناءً على مجموعة من القواعد:



يتيح لك ذلك التمرير لأعلى ولأسفل في السطر لمطابقة التنبؤ مع البيانات بشكل أفضل. إذا كان الثابت cغير موجود، سيمر الخط من خلال الأصل (0، 0) وسيكون لديك تطابق أضعف. ومن ثم، فإن التحيزات تسمح بتعلم المزيد والمزيد من الاختلافات في الأوزان. باختصار، تعني الاختلافات الإضافية أن التحيزات تضيف تمثيلًا أكثر ثراءً لمساحة الإدخال إلى الأوزان المستفادة من النموذج. لذلك، يتم حساب المعادلة النهائية للخلايا العصبية على النحو التالي:

$$\text{المخرج} = \sum (\text{المدخل} * \text{الوزن}) + \text{التحيز}$$



2. خوارزمية تعلم بيرسيبترون

تعلم بيرسبترون عملية بسيطة نسبيًا. هدفنا هو الحصول على مجموعة من الأوزان التي تصنف بدقة كل حالة في مجموعة التدريب الخاصة بنا. من أجل تدريب بيرسبترون، غالبًا ما نقوم بتغذية الشبكة ببيانات التدريب الخاصة بنا عدة مرات. في كل مرة ترى الشبكة مجموعة كاملة من بيانات التدريب، نقول إن دورة (epoch) قد مرت. الدورة هي معامل يحددها المستخدم قبل التدريب.
يمكن تلخيص شبيه الكود لخوارزمية تعلم Perceptron (الخوارزمية 1.3) كالاتي:

يتم "**التعلم** "الحقيقي في الخطوتين (2.ب) و (2.ج)، أولاً نقوم بتمرير متجه الميزة x_j عبر الشبكة، نحصل على ناتج المنتج الداخلي بأوزان w، ونحصل على الناتج y_j. بعد ذلك، يتم تمرير هذه القيمة للدالة الخطوية التي تُرجع 1إذا كانت $x > 0$ وإلا 0. الآن نحتاج إلى تحديث الوزن لدينا للتحرك في اتجاه أقرب إلى التصنيف الصحيح للبيانات. تتم إدارة هذه العملية عن طريق تحديث الوزن بموجب قانون دلتا الخطوة (2.ج)

يحدد التعبير $(d_j - y_j)$ ما إذا كان تصنيف المخرجات صحيحًا أم لا. إذا كان التصنيف صحيحًا، فسيكون هذا الاختلاف صفراً. خلاف ذلك، سيكون الاختلاف موجبًا أو سالبًا، مما يمنحنا اتجاهًا تظهر فيه الأوزان (مما يقربنا في النهاية من التصنيف الصحيح). ثم نضرب $(d_j - y_j)$ في xj، مما يقربنا من التصنيف الصحيح. قيمة α هي معدل التعلم (Learning rate) لدينا وتتحكم فيكبر (أو صغر) الخطوة. من المهم جدًا تعيين هذه القيمة بشكل صحيح. على الرغم من أن القيمة الكبيرة لـ α تسمح لنا باتخاذ خطوة في الاتجاه الصحيح، إلا أنه لا يزال بإمكاننا الوصول إلينا بسهولة من خلال التحسين المحلي أو العالمي. في المقابل، تسمح لنا كمية صغيرة من α باتخاذ خطوات صغيرة في الاتجاه الصحيح وتضمن أننا لا نتجاوز الحد الأمثل المحلي أو العالمي. ومع ذلك، قد تستغرق هذه الخطوات الصغيرة وقتًا طويلاً حتى يتقارب تعلمنا. أخيرًا، نضيف متجه الوزن السابق في الوقت t، $w_j(t)$ ، والذي يكمل عملية الانتقال نحو التصنيف الصحيح.

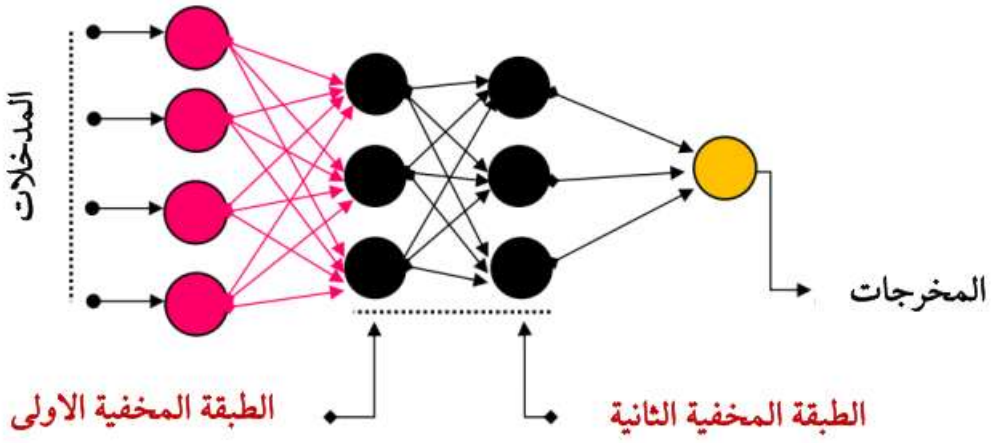
يسمح باستمرار عملية تعلم بيرسيبترون حتى يتم تصنيف جميع عينات التدريب بشكل صحيح أو الوصول إلى عدد محدد مسبقًا (محدد من قبل المستخدم) من الدورات. يتم ضمان الإنهاء إذا كانت αصغيرة بدرجة كافية وكانت بيانات التدريب قابلة للفصل خطيًا. بمعنى آخر، مع الافتراضات المناسبة، يمكن إظهار أن التعلم في بيرسيبترون سوف يتقارب مع الأوزان الصحيحة من خلال تكرار الخوارزمية الخاصة به. أي أن تعلم الشبكة سيؤدي إلى تقدير الأوزان التي تمكن الشبكة من إنتاج القيم الصحيحة عند المخرجات.

الخوارزمية 1.3	خوارزمية تعلم بيرسيبترون

1. ابدأ متجه الوزن w بقيم عشوائية صغيرة.
2. حتى يتقارب البيرسيبترون:
   أ. قم بعمل حلقة على كل متجه سمة x_j وقم بتسمية الفئة الفعلية d_j في مجموعة التدريب.
   ب. خذ x وقم بتمريرها عبر الشبكة وحساب قيمة الإخراج:
   $$y_j = f(w(t).x_j)$$
   ج. قم بتحديث الأوزان w:
   $$w: w_i(t+1) = w_i(t) + \alpha(d_j - y_j)x_{j,i}$$

3. بيرسيبترون متعدد الطبقات MLP (Multil Layer Perceptron)

كما هو مذكور، **فإن التقييد الرئيسي للشبكات العصبية بيرسيبترون هو عدم القدرة على تصنيف البيانات التي لا يمكن فصلها خطيًا.** يعد استخدام طبقة مخفية في هيكلية الشبكة بمثابة هروب من هذا التقييد. بمعنى آخر، لحل هذا التقييد، يمكن استخدام الطبقة المخفية بين طبقات الإدخال والإخراج. ومن الأمثلة على هذه الشبكات، والتي تعتبر أيضًا أساس التعلم العميق، شبكات بيرسيبترون متعدد الطبقات (**MultilLayer Perceptron**،(أو باختصار ، MLP والتي تسمى أيضًا الشبكات العصبية أمامية التغذية **(Feedforward Neural Network**تعد هذه الشبكات واحدة من أكثر الشبكات استخدامًا في التعلم العميق نظرًا لتوافقها مع مجموعة متنوعة من المشكلات. لأنه لا يوجد حد لإدخاله سواء كانت البيانات صورة أو نص أو فيديو.



شبكة عصبية ذات طبقتين مخفيتين

في MLP، تتدفق البيانات في الاتجاه الأمامي من طبقة الإدخال إلى الإخراج. في هذا النوع من الشبكات، بالانتقال من طبقة إلى أخرى، يتم حساب مجموع الاوزان لمجموعة الخلايا العصبية في الطبقة السابقة ونقلها إلى طبقة أخرى عن طريق تطبيق دالة تنشيط غير خطية. سبب تسمية التغذية الامامية **(Feed forward)** هو عدم وجود اتصال تغذية مرتدة يتم من خلاله اعتبار مخرجات النموذج مرة أخرى كمدخلات للنموذج (القيم تنتقل فقط من الإدخال إلى الطبقات المخفية ثم إلى الإخراج، ولا توجد قيم يتم إرجاعها إلى الطبقات السابقة، وعلى النقيض من ذلك، تسمح شبكة العودة بإرجاع القيم مرة أخرى). بمعنى آخر، الشبكة العصبية أمامية التغذية، تتدفق عمليات التنشيط في الشبكة دائمًا إلى الأمام عبر سلسلة من الطبقات. هذه الشبكة هي أيضًا شبكة متصلة بالكامل (**Fully Connected**،(لأن كل خلية من الخلايا العصبية للشبكة مترابطة بطريقة تستقبل مدخلات من جميع الخلايا العصبية في الطبقة السابقة وتنقل تنشيط مخرجاتها إلى جميع الخلايا العصبية في الطبقة التالية. يوضح الشكل 3-2مخططًا لشبكة عصبية مع طبقتين مخفيتين.

- الهدف من عملية تعلم الشبكة العصبية هو العثور على مجموعة من قيم الاوزان التي تجعل ناتج الشبكة العصبية مطابقا قدر الإمكان مع القيم المستهدفة الفعلي

الصعوبات التي تواجهنا أثناء تصميم وتدريب شبكة عصبية متعددة الطبقات MLP:

- ❖ تحديد عدد الطبقات المخفية المطلوبة.
- ❖ تحديد عدد الخلايا العصبية في كل طبقة مخفية.
- ❖ العثور على قيم مجموعة الأوزان المثلى.
- ❖ التقارب إلى الحل الأمثل في فترة زمنية معقولة (فترة تدريب قصيرة)
- ❖ التحقق من صحة نتائج الشبكة العصبية.