

السؤال الأول : اشرح بالتفصيل خوارزميات تدريب الشبكات العصبونية . (5 درجات)

الجواب : يعطى الطالب 5 درجات إذا ذكر جميع الخوارزميات مع الشرح كما يلي :

• خوارزمية (قاعدة) هيب (Hebb) – وجد دونالد هيب (Donald Hebb) في عام 1949 أن المشابك العصبية , الأوزان في الشبكات العصبونية , التي تتأثر أكثر من غيرها تصبح مثارة أي تزداد أوزانها بالمقارنة مع غيرها , حيث تعطى علاقة الخرج بالدخول على الشكل التالي :

$$\Delta W_{ij} = \gamma x_i y_j$$

حيث :

- معامل التعلم (سرعة التعلم) .
- Y_j – خرج (ناتج) العصبون j .
- x_i – دخل العصبون i .

تستخدم قاعدة هيب في عملية ضبط أوزان الذاكرة الترابطية (Associative Memory) حيث يكون $y = x$, وكذلك في بعض حالات التدريب بدون معلم .

إن هذه القاعدة تستخدم بكثرة و هي أثبتت فاعليتها , و لكن في الوقت الراهن توجد مجموعة من خوارزميات ضبط الأوزان , التي تعمل بشكل أفضل و القادرة على تدريب الشبكة بسرعة (زمن أقل) أكبر .

• خوارزمية (قاعدة) دلتا (Delta) – قام كل من بيرنارد وايدرو , مارسيان هوف (Bernard Widrow & Marcian Hoff) بوضع قاعدة لتعديل أوزان الشبكة العصبونية سميت بقاعدة دلتا , و تعتبر من أفضل الخوارزميات المستخدمة في ضبط أوزان الشبكات العصبونية أمامية التغذية ذات الطبقة الواحدة , حيث تعطى علاقة الخرج بالدخول على الشكل التالي :

$$\Delta W_{ij} = \gamma x_i (y_j^* - y_j)$$

حيث :

- معامل التعلم (سرعة التعلم) .
- $y_j^* - y_j$ – الفارق بين الناتج الهدف و الناتج الحالي .
- x_i – دخل العصبون i .

• خوارزمية (قاعدة) دلتا الموسعة (Extra Delta) – بالمقارنة مع قاعدة دلتا , تستخدم قاعدة دلتا الموسعة في الشبكات العصبونية متعددة الطبقات أمامية التغذية من المرتبة الثانية , حيث تعطى علاقة الخرج بالدخول على الشكل التالي :

$$\Delta W_{ij} = \gamma x_i \delta_j$$

حيث :

$$\delta_j = \begin{cases} f'_{aj}(x, w)(y_j^* - y_j) & ; \text{عندما يكون عصبون } j \text{ طبقة الخرج} \\ f'_{aj}(x, w) \sum_m \delta_m w_{jm} & ; \text{عند ذلك} \end{cases}$$

- مشتق تابع التحويل (التنشيط) .

• خوارزمية التنافس (Competitive) – في هذا النوع من خوارزميات تعديل الأوزان تعتبر كل العصبونات مثيلة بعضها البعض , ضمن الطبقة الواحدة , و العصبون الأكثر تأثراً , العصبون الرابع , فقط يعدل وزنه (Winner takes all) .

أوزان بقية العصبونات تبقى ثابتة , و يمكن أن تتناقص في بعض الحالات , و عملية تعديل (ضبط) الأوزان تنتهي عندما يصبح وزن العصبون الرابع مساوياً لمجموع أوزان بقية العصبونات الطبقة نفسها .

• الخوارزمية العشوائية (Stochastic) – في هذا النوع من خوارزميات تعديل الأوزان لا يخضع لقوانين ثابتة و محددة بل تتم عمليات تعديل الأوزان بشكل شبه عشوائي مع الحفاظ على أوزان العصبونات الأفضل .

يستخدم هذا النوع من الخوارزميات في كل من آلة كاوتشي (Cauchy) و بولتزمان (Boltzmann) .

• خوارزمية إنقاص النسب (Gradient Descent) – عندما تكون توابع التحويل (التنشيط) قابلة للاشتقاق فإنه من المفيد استخدام هذا النوع من خوارزميات تعديل الأوزان , حيث تعطى علاقة الخرج بالدخول على الشكل التالي :

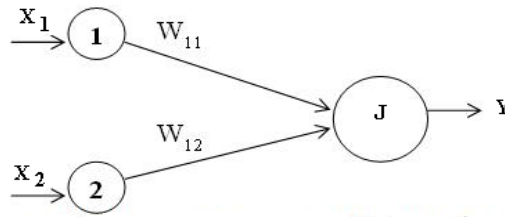
$$\Delta W_{ij} = - \gamma \frac{\partial I}{\partial w_{ij}}$$

حيث :

تعتبر كل من قاعدة ضبط الأوزان دلنا , قاعدة **Widrow , Hoff** , وطريقة التدریب " الانتشار العكسي " من أشكال خوارزمية إنقاص النسب .

ملاحظة : تحذف علامة واحدة على كل خوارزمية لم تذكر أو لم تشرح .

السؤال الثاني : لنفرض أن لدينا شبكة (Perceptron) كما في الشكل و نريد تدريبها : (5 درجات)



ونريد تعليمها كيف تطبق الدالة المنطقية OR , وقاعدة التحويل (تابع التحويل) هي :

$$\begin{aligned} \text{if } S_j > 0 \text{ then } Y_j &= 1 \\ \text{if } S_j \leq 0 \text{ then } Y_j &= 0 \end{aligned}$$

ولیکن معدل (معامل) التعلم معطى بالقيمة $C=0.5$, و لیکن لهذه الشبكة قيم أوزان مختلفة على أن تكون في المدى من -0.5 إلى 0.5 و لنفرض القيم التالية : $w_{11}=0.1$, $w_{12}=0.3$.

الجواب : يعطى الطالب 5 درجات إذا تضمن جواب الطالب ما يلي :

- نحسب خرج الجامع (للحالة الأولى) : $S = 0.0 \times 0.1 + 0.0 \times 0.3 = 0.0$
- نحدد خرج تابع التفعيل : $S = 0.0 \leftarrow S \leq 0.0$ إذا $Y = 0$.

نلاحظ أن خرج الشبكة صحيح وهذا يدل على أن أوزان الشبكة ليست بحاجة إلى تعديل .

نحسب خرج الجامع (للحالة الثانية) مستخدمين نفس الأوزان السابقة لأنه لم يتم تعديلها لعدم الحاجة لذلك :

- نحسب خرج الجامع : $S = 0.0 \times 0.1 + 1.0 \times 0.3 = 0.3$
- نحدد خرج تابع التفعيل : $S = 0.3 \leftarrow S > 0.0$ إذا $Y = 1$.

نلاحظ أن خرج الشبكة صحيح وهذا يدل على أن أوزان الشبكة ليست بحاجة إلى تعديل .

نحسب خرج الجامع (للحالة الثالثة) مستخدمين الأوزان الجديدة

- نحسب خرج الجامع : $S = 1.0 \times 0.1 + 0.0 \times (0.3) = 0.1$
- نحدد خرج تابع التفعيل : $S = 0.1 \leftarrow S > 0.0$ إذا $Y = 1$.

نلاحظ أن خرج الشبكة صحيح وهذا يدل على أن أوزان الشبكة ليست بحاجة إلى تعديل .

نحسب خرج الجامع (للحالة الرابعة) مستخدمين الأوزان الجديدة

- نحسب خرج الجامع : $S = 1.0 \times 0.1 + 1.0 \times (0.3) = 0.4$
- نحدد خرج تابع التفعيل : $S = 0.4 \leftarrow Y = 1$.

نلاحظ أن خرج الشبكة صحيح وهذا يدل على أن أوزان الشبكة ليست بحاجة إلى تعديل .

السؤال الثالث : أوجد حلا للسؤال السابق باستخدام خوارزمية الانتشار العكسي (الخلفي). (5 درجات)

الجواب : يعطى الطالب 5 درجات إذا تضمن جواب الطالب ما يلي :

لحل السؤال يجب استخدام تابع التفعيل الأسّي عوضاً عن تابع تفعيل العتبة , و لتكن الأوزان الابتدائية للشبكة كالاتي : $w_{11}=0.5$, $w_{12}=0.3$.

نأخذ إحدى الحالات الأربعة لبوابة OR , و لتكن الحالة الرابعة و نطبق عليها خوارزمية الانتشار العكسي كما يلي :

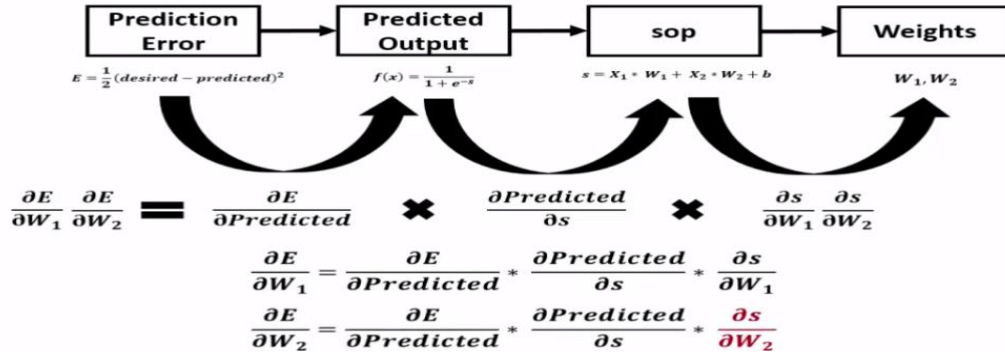
نحسب خرج الجامع (للحالة الرابعة) مستخدمين الأوزان الجديدة

• نحسب خرج الجامع : $S = 1.0 \times 0.5 + 1.0 \times (0.3) = 0.8$

• نحدد خرج تابع التفعيل : $Y = S / (1 + e^{-S})$ ← $E = Y^* - Y$

نلاحظ أن هناك خطأ خرج الشبكة غير صحيح وهذا يدل على أن أوزان الشبكة بحاجة إلى تعديل و عليه نطبق خوارزمية الانتشار العكسي كما يلي :

Chain Rule



$$\frac{\partial E}{\partial \text{Predicted}} = \text{predicted} - \text{desired}$$

$$\frac{\partial \text{Predicted}}{\partial s} = \frac{1}{1 + e^{-s}} \left(1 - \frac{1}{1 + e^{-s}}\right)$$

$$\frac{\partial s}{\partial W_i} = X_i$$

$$\frac{\partial E}{\partial W_1} = \frac{\partial E}{\partial \text{Predicted}} * \frac{\partial \text{Predicted}}{\partial s} * \frac{\partial s}{\partial W_1}$$

$$\frac{\partial E}{\partial W_2} = \frac{\partial E}{\partial \text{Predicted}} * \frac{\partial \text{Predicted}}{\partial s} * \frac{\partial s}{\partial W_2}$$

بالتعويض نوجد القيم المطلوبة .

نحسب الأوزان الجديدة وفق الآتي :

$$W_{inew} = W_{iold} - \eta * \frac{\partial E}{\partial W_i}$$

انتهى السلم