الجمهورية الجزائرية الديمقراطية الشعبية
وزارة التعليم العالي
والبحث العلمي

جامعة قاصدي مرباح ورقلة
قسم الرياضيات وعلوم المادة
كلية الرياضيات
تخصص : احصاء و احتمالات

مذكرة مقدمة لقسم الرياضيات لإستكمال شهادة الماستر 02

من إعداد:

رحمون سيد علي

مذكرة تحت عنوان:

تقدير الإنحدار الخطي المتعدد بإستعمال الشبكات العصبية

اعضاء لجنة المناقشة:

كارك محمد	جامعة قاصدي مرباح ورقلة	رئيسا
عقون رشيد	جامعة قاصدي مرباح ورقلة	مشرفا
عباسي حسين	جامعة قاصدي مرباح ورقلة	مناقشا

السنة الجامعية 2023-2024

إخلاص

إلى جدّتي،

إلى والديَّ،

إلى إخوتي،

وإلى جميع أفراد عائلتي،

إلى أصدقائي الأعزاء محمد العيد، نور الدين، عبد المنعم، نور الدين، فاضل، محمد، أيمن، أحمد، محمد عصام،

عبد السميع، عبد الله، عبد السلام، علي،

إلى صديقتي وفاء،

أود أن أشكر مشرفي:

الدكتور عقون رشيد ,

وجميع أعضاء فصلي وبعض أساتذتي،

إلى كل من عرقلني،إلى كل محبط،

إلى أعداء النجاح،

وأخيراً، لكل من ساعدني بأي شكل من الأشكال،

إلى من يسعهم قلبي ولم تسعهم هذه الورقة.

شكر

بدأت بأكثر من يد وقاسيت أكثر من هم وعانيت الكثير من الصعوبات وها أنا اليوم والحمد لله أطوي سهر الليالي وخلاصة مشواري بين دفتي هذا العمل المتواضع الذي أهديه:

إلى من بلغ الرسالة وأدى الأمانة ونصح الأمة الى نبي الرحمة رسولنا الكريم عليه أفضل الصلاة وأزكى التسليم صلو على رسول الله صلى الله عليه وسلم.

إلى من قال الله عز وجل في حقهما:

(وقضى ربك ألا تعبدوا إلا إياه وبالوالدين إحسانا).

إلى بهجة القلب وهبة الرب وكمال الود وصفاء الحب

الى التي حملت الحياة بين يديها بريقا وشعاعا لدربي.

الى التي اهدتني رضاها ولم تبخل عني بدعواتها

(أمي الغالية حفضها الله).

إلى مصدر فخري

أبي.

إلى إخوتي وجميع أفراد عائلتي

إلى كل من اجتمعت بهم على محبة الله وطاعته وتوحدنا على دعوته.

إلى كل من شاطرني ثمرة جهدي.

الى من يسعهم قلبي ولم تسعهم هذه الورقة.

قائمة الرموز

- Y هي مصفوفة النواتج بأبعاد $m \times 1$.
- X هي مصفوفة المتغيرات المستقلة بأبعاد $m \times n$.
- W هي مصفوفة الأوزان بأبعاد $n \times 1$.
- $\hat{y}_i$ هي القيمة المتوقعة .
- n هو عدد العينات في بيانات التدريب.
- ∇: هو التدرج
- $m_{(\theta)}^{[t]}$: هو تقدير اللحظة الأولى (أي متوسط الحركة).
- $v_{(\theta)}^{[t]}$: هو تقدير اللحظة الثانية (أي متوسط الحركة المربعة).
- β_1 و β_2: هما معاملات التحكم في تلاشي اللحظات.
- $(\odot)$: يشير إلى الضرب التكويني للمصفوفات (عنصر بعنصر).
- $\hat{m}_{(\theta)}^{[t]}$ و $\hat{v}_{(\theta)}^{[t]}$: هي تصحيحات التحيز لـ $m_{(\theta)}^{[t]}$ و $v_{(\theta)}^{[t]}$ على التوالي.
- η: هو معدل التعلم.
- ϵ: هو مصطلح التحسين لتجنب القسمة على صفر.
- g: دالة التنشيط.
- $J(W, b)$: هي دالة الخسارة الأصلية.
- λ: هو معامل التسوية.
- MSE: متوسط الخطأ التربيعي.
- MAE: الخطأ المطلق المتوسط.
- $RMSE$: الجذر التربيعي للخطأ التربيعي الوسطي.
- R^2: معامل الارتباط.
- ^{T} يدل على العملية المتجانسة .(Transpose)
- $^{-1}$ يدل على العملية العكسية .(Inverse)
- $J(W, b)$: دالة الخسارة.

- W: الأوزان في الشبكة العصبية.
- b: الانحيازات في الشبكة العصبية.
- $y^{(i)}$: القيمة الحقيقية للعينة i.
- $A^{[L](i)}$: الإخراج للطبقة الأخيرة L للعينة i.
- $E^{[L](i)}$: خطأ الطبقة الأخيرة L للعينة i.
- $Z^{[L](i)}$: مجموع المدخلات المرجحة للطبقة الأخيرة L للعينة i.
- $W^{[L]}$: الأوزان بين الطبقة الأخيرة L والطبقة السابقة لها $L-1$.
- الطبقات الوسطى $l = 2, \ldots, L-1$
- الطبقة الأخيرة L

قائمة الأشكال

1.1 مكونات الشبكة العصبية الاصطناعية 5
2.1 منحنى لدالة التنشيط ريلو وسيغمويد 6

1.2 الإنتشار العكسي 16
2.2 خوارزمية التحسين Adam 17
3.2 تقارب دالة الخطأ لبيانات التدريب 20
4.2 شرح طريقة عمل الشبكة العصبية 26

1.3 موقع البئر العمودي 27
2.3 تثبيت المكتبات وتحميلها 28
3.3 الحصول على نتائج متكررة 29
4.3 تحميل البيانات 29
5.3 عملية الإستيفاء الخطي 29
6.3 تحديد المتغيرات المستقلة والمتغير التابع 30
7.3 تقسيم البيانات الى بيانات تدريب وأخرى للإختبار 30
8.3 تطبيع البيانات 30
9.3 إنشاء نموذج شبكة عصبية 30
10.3 ترجمة النموذج 31
11.3 تدريب النموذج 31
12.3 مسار التطور أثناء التدريب 31
13.3 تقييم النموذج 31
14.3 نتائج التقييم 32
15.3 إجراء تنبؤات 32
16.3 حساب مقاييس الأداء 32
17.3 رسم القيم الفعلية مقابل القيم المتنبأة ومسار التطور خلال التدريب. 33
18.3 رسم القيم الفعلية مقابل القيم المتنبأ بها 33
19.3 رسم خط الإنحدار 33
20.3 رسم مسار التطور أثناء التدريب 34
21.3 رسم القيم الفعلية مقابل القيم المتنبأ بها بصيغة أفضل 34
22.3 رسم القيم الفعلية مقابل القيم المتنبأ بها بصيغة أفضل 34
23.3 حساب ورسم البواقي 35
24.3 رسم المدرج التكراري للبواقي 36
25.3 رسم البواقي مقابل القيم المتوقعة 37

المحتويات

إخلاص 1

شكر 2

الفهرس 6

1 الأساسيات النظرية 2
1.1 الانحدار الخطي 2
2.1 الشبكات العصبية: 4
3.1 الشبكات العصبية في مشاكل الانحدار [9] 6

2 تقدير الانحدار الخطي باستخدام الشبكات العصبية 8
1.2 مدخل 8
2.2 مقدمة 8
3.2 الجانب النظري 8
4.2 دوال التنشيط في الشبكات العصبية 9
5.2 التدريب والانتشار الأمامي والعكسي 12
6.2 تحسين النموذج [20] 16
7.2 تقنيات منع الإفراط في التعلم [16],[12] 23
8.2 تقييم النموذج 25

3 برمجة شبكة عصبية لتقدير الانحدار الخطي المتعدد بإستخدام لغة البرمجة R 27

مقدمة

يلعب تحليل الانحدار دورًا كبيرًا في دراسة الظواهر المهمة في مختلف مجالات الحياة، وذلك من خلال تحليل هذه الظواهر والتنبؤ بنتائجها المستقبلية. يعتبر نموذج الانحدار وسيلة إحصائية قوية تعبر عن العلاقة بين المتغيرات، حيث يمثل المتغير التابع الظاهرة المدروسة، وتُستخدم المتغيرات المستقلة لتفسير تلك الظاهرة. يهدف تحليل الانحدار إلى تقدير دالة الانحدار باستخدام طرق وسيطية وغير وسيطية.

في هذه المذكرة، يتم التركيز على طريقة تقدير الانحدار الخطي المتعدد باستخدام الشبكات العصبية، وهي إحدى تقنيات التعلم الآلي التي تتميز بقدرتها على معالجة البيانات المعقدة واكتشاف الأنماط المخفية فيها. تحظى الشبكات العصبية بمرونة عالية مقارنةً بالطرق التقليدية مثل طريقة المربعات الصغرى، التي تعد واحدة من أكثر الطرق شيوعًا لتقدير معلمات نموذج الانحدار.

تتميز الشبكات العصبية بقدرتها على تقديم نماذج دقيقة وقابلة للتعميم من خلال تدريبها على مجموعات بيانات كبيرة ومتنوعة. يتم في هذه المذكرة استخدام لغة البرمجة R لتطبيق الشبكات العصبية على تقدير الانحدار الخطي المتعدد، مما يتيح استكشاف العلاقات المعقدة بين المتغيرات وتحسين دقة التنبؤات.

من خلال هذا النهج، نسعى إلى تقديم رؤية جديدة ومبتكرة لتحليل البيانات باستخدام تقنيات حديثة، مما يسهم في تعزيز القدرة التنبؤية للنماذج وتقديم حلول فعالة للتحديات التي تواجهها مختلف المجالات.

تتألف هذه المذكرة من ثلاثة فصول، موزعة كالتالي:

- **الفصل الأول: الأساسيات النظرية**
- **الفصل الثاني: تقدير الانحدار الخطي المتعدد باستخدام الشبكات العصبية**
- **الفصل الثالث: برمجة شبكة عصبية لتقدير الانحدار الخطي المتعدد باستخدام لغة البرمجة R**

باب 1

الأساسيات النظرية

1.1 الانحدار الخطي

1.1.1 تعريف:

الانحدار الخطي هو تقنية إحصائية تُستخدم لدراسة العلاقة بين متغير تابع وواحد أو أكثر من المتغيرات المستقلة. الهدف هو نمذجة العلاقة بين هذه المتغيرات بخط مستقيم (للانحدار الخطي البسيط) أو بسطح (للانحدار الخطي المتعدد) يمثل أفضل تقدير للعلاقة بين هذه المتغيرات.

2.1.1 أهمية الانحدار الخطي:

الانحدار الخطي مهم في البحث العلمي وتحليل البيانات لأنه يسمح بفهم كيفية تأثير المتغيرات المستقلة على المتغير التابع وتقدير هذه العلاقة كميًا.

3.1.1 الفرق بين الانحدار الخطي البسيط والمتعدد:

الانحدار الخطي البسيط ينطوي على متغير مستقل واحد فقط لتقدير قيم المتغير التابع

$$y = \beta_0 + \beta_1 x + \epsilon$$

بينما الانحدار الخطي المتعدد يستخدم اثنين أو أكثر من المتغيرات المستقلة للتنبؤ بالمتغير التابع.

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n + \epsilon$$

$$Y = X\beta + \epsilon \tag{1}$$

حيث:

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \quad X = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1p} \\ 1 & x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix}, \quad \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix}, \quad \epsilon = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{bmatrix}$$

4.1.1 أنواع الانحدار الخطي المتعدد:

هناك ثلاثة نماذج انحدار متعدد رئيسية هي:

- الانحدار العياري أو القياسي .(standard)
- الانحدار الهرمي .(Hierarchical)

- الانحدار المتدرج .(stepwise)

هذه النماذج تختلف من جهتين:

- الأولى في معالجة الاختلافات المتداخلة بسبب ارتباط المتغيرات المستقلة.
- الثانية في ترتيب إدخال المتغيرات المستقلة في المعادلة.

على ضوء هذين الاختلافين فإن نماذج الانحدار يمكن تقسيمه إلى الأقسام التالية:

1. الانحدار المعياري: في هذه الطريقة ندخل المتغيرات المستقلة في معادلة الانحدار دفعة واحدة لنحصل على المعادلة التي تصف العلاقة بين كل المتغيرات المستقلة والمتغير التابع مرة واحدة دون مناقشة هل كل المتغيرات المستقلة يجب أن تدخل في المعادلة أم لا؟ ولا نتعرض لمناقشة هل المتغيرات المستقلة مرتبطة بعضها ببعض أم مستقلة.

2. الانحدار الهرمي: في الانحدار الهرمي تدخل المتغيرات المستقلة في المعادلة المقترحة تباعاً ونحدد ترتيب دخول هذه المتغيرات في المعادلة المقترحة على أساس إحصائي نظري.

3. الانحدار التدريجي: في نموذج الانحدار المتدرج، عدد المتغيرات المستقلة المدخلة في النموذج وكذلك ترتيب إدخالها يحدد من خلال معيار إحصائي يتم الوصول إليه عن طريق إجراء الانحدار المتدرج.

5.1.1 خواص الانحدار الخطي المتعدد

الانحدار الخطي المتعدد هو تقنية إحصائية تُستخدم لنمذجة العلاقة بين متغير تابع واحد (المتغير الذي نحاول التنبؤ به) وعدة متغيرات مستقلة (المتغيرات التي نستخدمها للتنبؤ). فيما يلي بعض الخواص الأساسية للانحدار الخطي المتعدد:

- **الخطية**: يعتمد الانحدار الخطي المتعدد على فرضية أن العلاقة بين المتغير التابع وكل من المتغيرات المستقلة خطية، أي يمكن تمثيلها بمعادلة خط مستقيم في شكل:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_n X_n + \epsilon$$

حيث Y هو المتغير التابع، $X_1, X_2, \ldots, X_n$ هي المتغيرات المستقلة، $\beta_0, \beta_1, \ldots, \beta_n$ هي المعاملات التي تمثل تأثير المتغيرات المستقلة على المتغير التابع، وϵ هو الخطأ العشوائي.

- **الافتراضات**:
  - **الاستقلالية**: يجب أن تكون ملاحظات البيانات مستقلة عن بعضها البعض.
  - **التوزيع الطبيعي للأخطاء**: يجب أن تتبع الأخطاء (الفرق بين القيم الفعلية والقيم المتوقعة) توزيعاً طبيعياً.
  - **التجانس**: يجب أن يكون تباين الأخطاء ثابتاً عبر جميع مستويات المتغير المستقل.
  - **عدم وجود التعدد الخطي الكامل**: لا ينبغي أن يكون هناك علاقة خطية كاملة بين أي اثنين أو أكثر من المتغيرات المستقلة.

- **تفسير المعاملات**: معامل الانحدار β_i يمثل التغير في المتغير التابع Y نتيجة تغير المتغير المستقل X_i بوحدة واحدة، مع بقاء المتغيرات المستقلة الأخرى ثابتة.

- **التحقق من النموذج**: يمكن استخدام مقاييس مثل R^2 لتقييم جودة النموذج ومدى تفسيره للتباين في البيانات. تحليل البواقي يمكن أن يساعد في التحقق من الافتراضات مثل التجانس والتوزيع الطبيعي للأخطاء.

- **التنبؤ**: يمكن استخدام النموذج المناسب للتنبؤ بالقيم المستقبلية للمتغير التابع بناءً على القيم الجديدة للمتغيرات المستقلة.

- **التحديات**:
  - **التعدد الخطي**: إذا كانت هناك علاقة خطية قوية بين المتغيرات المستقلة، يمكن أن يؤثر ذلك على دقة تقدير المعاملات.
  - **تأثير النقاط الشاذة**: النقاط التي تكون بعيدة جداً عن بقية البيانات يمكن أن تؤثر بشكل كبير على تقديرات النموذج.
  - **الافتراضات**: إذا لم تُحقق الافتراضات المذكورة أعلاه، فإن نتائج النموذج قد تكون مضللة.

6.1.1 التقدير:

التقدير هو عملية استخدام البيانات الملاحظة لتحديد معلمات نموذج إحصائي. الهدف من التقدير هو الحصول على تقديرات للمعلمات غير المعروفة التي تصف العلاقة بين المتغيرات في النموذج. هناك عدة طرق وأدوات مستخدمة في التقدير، ومنها:

- **طريقة المربعات الصغرى:** يستخدم نموذج الانحدار الخطي المتعدد طريقة المربعات الصغرى العادية (OLS) كأسلوب أساسي لتقدير المعاملات، حيث يهدف إلى تقليل مجموع مربعات الفروق بين القيم الملاحظة للمتغير التابع والتنبؤات التي يقدمها النموذج. توضح الدراسة التي قام بها عبد الفتاح مصطفى محمد في كلية العلوم، جامعة مصر، أهمية وكفاءة استخدام OLS في التعامل مع البيانات المتعددة المتغيرات، مما يُمكّن الباحثين من تحليل العلاقات المعقدة بدقة أكبر [5].
- **معادلة تقدير طريقة مربعات** معادلة تقدير طريقة مربعات (OLS) تعطى بالصيغة التالية:

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y} \tag{2}$$

حيث:

- $\hat{\boldsymbol{\beta}}$ هو متجه التقديرات لمعاملات الانحدار.
- $\mathbf{X}$ هي مصفوفة التصميم التي تحتوي على القيم المرصودة للمتغيرات المستقلة وعمود من الواحدات للقطع الصفري.
- $\mathbf{y}$ هو متجه القيم المرصودة للمتغير التابع.
- ^{T} يدل على العملية المتجانسة .(Transpose)
- $^{-1}$ يدل على العملية العكسية .(Inverse)

تحليل التباين:
تحليل التباين (ANOVA) هو طريقة إحصائية تُستخدم لمقارنة متوسطات مجموعتين أو أكثر لمعرفة ما إذا كانت هناك فروق ذات دلالة إحصائية بينها. [7].

$$SS_{\text{الكلي}} = SS_{\text{الانحدار}} + SS_{\text{الخطأ}} \tag{3}$$

حيث:

- $SS_{\text{الكلي}}$ هو مجموع مربعات الانحرافات حول المتوسط العام، ويمثل التباين الكلي في البيانات.
- $SS_{\text{الانحدار}}$ هو مجموع مربعات الانحرافات للقيم المتوقعة حول المتوسط العام، ويمثل التباين الذي يفسره النموذج.
- $SS_{\text{الخطأ}}$ هو مجموع مربعات الانحرافات للقيم الملاحظة حول القيم المتوقعة، ويمثل التباين الناتج عن الخطأ أو العشوائية.

2.1 الشبكات العصبية:

مدخل

الشبكات العصبية هي أنظمة حسابية مستوحاة من الشبكات العصبية البيولوجية التي تشكل أدمغة الكائنات الحية. تاريخيًا، نشأت فكرة الشبكات العصبية من الرغبة في فهم كيف يستطيع الدماغ البشري أداء مهام معقدة بكفاءة. الأعمال الأولية في هذا المجال يمكن تتبعها إلى الأربعينيات، لكن الاهتمام بهذه التقنية شهد تجددًا قويًا في الثمانينات والتسعينات بفضل التطورات في الحوسبة والخوارزميات.

ونجد أيضا تعريف الشبكة العصبية الاصطناعية ينص على أنَّها: تقنيات حاسوبية لمعالجة المعلومات، مكونة من عدد كبير جدًا من عناصر المعالجة المترابطة فيما بينها ذات طبيعة ديناميكية وظيفية، وهي التقسيم المتوازي للشبكة. تتكون الشبكة العصبية من عدد من وحدات المعالجة المتداخلة والمتجانسة، كل وحدة بمفردها أداة حاسوبية يمكن نمذجة سلوكها بمعادلات رياضية بسيطة [1].

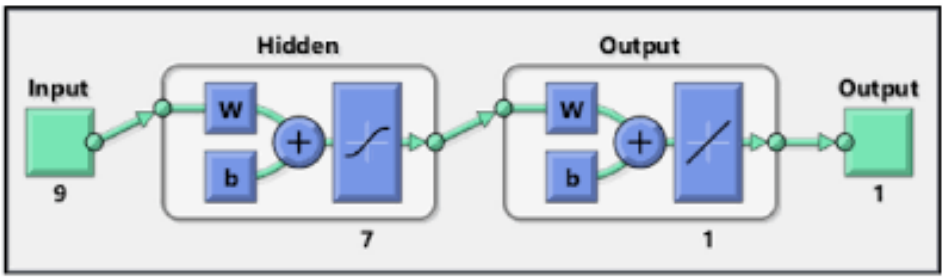
وعرَّفها أحد الباحثين بأنها: محاولة لإيجاد تقنيات حاسوبية تُصمَّم لمحاكاة الطريقة التي يؤدي بها الدماغ البشري مهمة معينة، وتتشابه معه في أنها تكتسب المعرفة بالتدريب المستمر وخزن هذه المعرفة، ومن خلال ذلك يتم الوصول إلى الحل الأمثل للمسألة قيد الدراسة، إذ تكتسب الشبكة المعرفة بالتجربة من خلال الاطلاع على البيانات التاريخية والاستفادة من التجارب السابقة [6].

كما عُرِّفت بأنها نموذج رياضي أو حسابي مبني على الشبكات العصبية الحيوية، أي أنها تحاكي النظام العصبي الحيوي، وتتكون من مجموعة من الخلايا العصبية والعمليات المعلوماتية المترابطة، وهي قابلة للتكيف اعتمادًا على المعلومات الخارجية أو الداخلية التي تجري خلال الشبكة أثناء مرحلة التعلم [4].

1.2.1 الشبكات العصبية ذات التغذية الخلفية:

في هذه الشبكات من الممكن أن تُعاد تغذية كل أو بعض الإشارات الخارجة من الشبكة، وتحويل اتجاهها لتصبح إشارة داخلة أيضًا (إعادة تغذية)، وبذلك فإن الإشارة الخارجة من أي خلية تعتمد على الإشارات الداخلة إليها، فضلاً عن الإشارات الخارجة منها في وقت سابق. الشكل الآتي يوضح هذا النوع من الشبكات، كما أن خوارزمية التدريب هي خوارزمية الانتشار العكسي نفسها [3].

2.2.1 هيكل الشبكة العصبية: الوحدات العصبية، الطبقات، دوال التنشيط



شكل 1.1: مكونات الشبكة العصبية الاصطناعية

مكونات الشبكة العصبية الاصطناعية: كما أن للإنسان وحدات إدخال توصله بالعالم الخارجي، وهي حواسه الخمس، كذلك فالشبكات العصبية لها مدخلات تمثل في وحدات المعالجة التي تتم فيها العمليات الحسابية التي تضبط بها الأوزان ونحصل عن طريقها على ردة الفعل لكل مدخل من المدخلات للشبكة. فوحدات الإدخال تكوّن طبقة تسمى طبقة المدخلات، ووحدات المعالجة تكوّن طبقة المعالجة، والتي تقوم بإخراج نتائج الشبكة وتدعى طبقة المخرجات. وبين كل طبقة من هذه الطبقات هناك طبقة من الوصلات البينية التي تربط كل طبقة بالطبقة التي تليها [2].
ونوضحها فيما يلي[3]:

- **الوحدات العصبية (النيورونات):** البنية الأساسية للشبكة العصبية، حيث تستقبل كل وحدة إشارات من الوحدات الأخرى، تعالج هذه الإشارات، ثم تنقل الإشارة المعالجة إلى الوحدات التالية.

- **الطبقات:** تُقسم الشبكات العصبية إلى طبقات، بدءًا من طبقة الإدخال، مرورًا بطبقات مخفية واحدة أو أكثر، وانتهاءً بطبقة الإخراج. تعالج كل طبقة جزءًا من المهمة وتنقل النتائج إلى الطبقة التالية.

- **المدخلات:** تتكون طبقة المدخلات من مجموعة من الوحدات تعمل على توزيع القيم المدخلة إليها من الوسط الخارجي إلى الطبقة الخفية التي تمر عبر الارتباطات الموجودة بينها. وقد تتألف هذه الشبكة حسب تركيبتها من وحدة معالجة أو أكثر، فمدخلات الشبكة يمكن أن تكون بصورة بيانات خام، بيانات ثنائية أو بيانات متصلة، كما يمكن أن تكون من وحدات معالجة أخرى. بذلك فإن مدخلات الشبكة لا يتم فيها معالجة حسابية، بل تقوم بنقل البيانات إلى الطبقة الخفية عبر الوصلات البينية أو مباشرة إلى طبقة المخرجات إذا كانت الشبكة لا تحوي طبقة خفية .

- **المخرجات:** وهي الطبقة الأخيرة التي تتكون من وحدات معالجة تتبع الطبقات السابقة وتعطي الإخراج الحقيقي. كما يمكن أن تحوي هذه الطبقة على وحدة معالجة واحدة أو أكثر وهذا راجع للبنية المعمارية للشبكة العصبية. تقوم هذه الطبقة باستقبال الإشارات القادمة من طبقة الإدخال وطبقة الخفية، وإعادة إرسالها بعد إجراء المعالجات عليها أو إعادة هذه المخرجات للشبكة عندما لا تستطيع معالجة بياناتها. فالمخرجات تمثل حل المشكلة حيث تعطي الشبكة قيما عددية للمخرجات وغالبا ما تحوي الشبكة على طبقة إخراج واحدة [?].

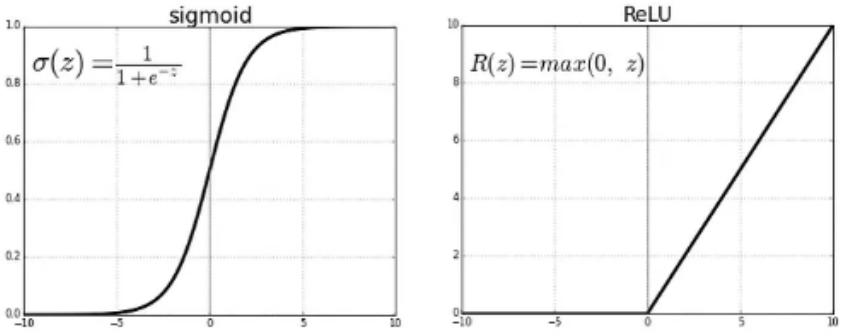
- **الطبقات الخفية:** هي طبقة موجودة بين طبقة المدخلات وطبقة المخرجات، حيث أنه يمكن أن تحتوي الشبكة على طبقة خفية واحدة أو أكثر من طبقة خفية. تقوم الطبقة الخفية باستقبال الإشارات القادمة إليها من طبقة المدخلات عبر الوصلات البينية من أجل معالجتها وإجراء اللازم عليها ومن ثم تقوم هي بدورها بإرسالها عبر الوصلات إلى طبقة المخرجات [2].

- **دوال التنشيط:** دوال التنشيط هي دوال رياضية تُستخدم في الشبكات العصبية الاصطناعية لتحديد ناتج كل عصبون. الهدف الرئيسي من استخدام دوال التنشيط هو إدخال اللاخطية إلى نموذج الشبكة العصبية، مما يسمح لها بتعلم الأنماط والعلاقات المعقدة في البيانات.

تعمل دوال التنشيط على تحويل المدخلات الخطية إلى مخرجات غير خطية، مما يساعد الشبكة على تعلم مختلف الأنماط بدقة عالية. تُستخدم دوال التنشيط في الطبقات المخفية والطبقة النهائية للشبكة العصبية لتحديد كيفية انتقال الإشارات عبر العصبونات.

هناك عدة أنواع من دوال التنشيط، كل منها لها خصائصها واستخداماتها المناسبة في تطبيقات معينة. من بين الدوال الشهيرة نذكر دالة سيغمويد، دالة تان-إتش، دالة ريلو، ودالة ليكي ريليو.

بفضل هذه الدوال، يمكن للشبكات العصبية التعامل مع المشاكل المعقدة والمتعددة الأبعاد في مجالات مثل التصنيف، التنبؤ، ومعالجة الصور والنصوص.[19]



شكل 2.1: منحنى لدالة التنشيط ريلو وسيغمويد

3.1 الشبكات العصبية في مشاكل الانحدار [9]

في مشاكل الانحدار، الهدف هو تقدير قيمة متغير مستمر بناءً على واحد أو أكثر من المتغيرات المستقلة. الشبكات العصبية يمكن أن تتعامل مع هذا النوع من المهام بكفاءة عالية من خلال تعديل وزنها وثوابتها الانحيازية بناءً على بيانات التدريب. ونستطيع أن نقول أيضا، تُستخدم الشبكات العصبية لتوقع قيمة مستمرة استنادًا إلى مجموعة من المدخلات، مثل توقع أسعار المنازل استنادًا إلى المواصفات والموقع. تعد الشبكات العصبية أداة قوية في هذا المجال بفضل قدرتها على التعامل مع كميات كبيرة من البيانات وتعقيدها، واكتشاف العلاقات الخطية وغير الخطية بين المتغيرات.

1.3.1 كيفية عملها في الانحدار:

1. **تهيئة الشبكة:** تُحدد طبقات الشبكة وعدد النيورونات في كل طبقة، بالإضافة إلى اختيار دالة التنشيط المناسبة لطبقات الشبكة الخفية.

2. **اختيار دالة الخسارة:** في مشاكل الانحدار، دالة الخسارة الأكثر شيوعًا هي متوسط الخطأ التربيعي ،(MSE) والتي تقيس الفرق بين القيم المتوقعة من الشبكة والقيم الفعلية من بيانات التدريب.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - A^{[L](i)})^2$$

3. **التدريب والتعديل:** يتم استخدام خوارزميات التحسين، مثل الهبوط التدريجي لتعديل الأوزان والثوابت الانحيازية في الشبكة بناءً على الخطأ المحسوب بواسطة دالة الخسارة. الهدف هو تقليل دالة الخسارة وبالتالي زيادة دقة توقعات الشبكة.

4. **التقييم والتحسين:** بعد التدريب، يتم تقييم أداء الشبكة على بيانات لم تُستخدم في عملية التدريب (عادة ما تُعرف ببيانات الاختبار) لتقييم دقتها وتعميم قدرتها على بيانات جديدة.

عند تقييم أداء نماذج الانحدار، هناك عدة معايير تقييم رئيسية يمكن استخدامها:

1. **متوسط الخطأ المطلق (MAE):** يقيس متوسط الفروق المطلقة بين القيم المتوقعة والقيم الفعلية.
2. **متوسط الخطأ التربيعي (MSE):** يقيس متوسط الفروق التربيعية بين القيم المتوقعة والقيم الفعلية، مما يعطي وزناً أكبر للأخطاء الكبيرة.
3. **الجذر التربيعي لمتوسط الخطأ التربيعي (RMSE):** يعتبر الجذر التربيعي لـ MSE ويقدم الأخطاء بنفس وحدات المتغير التابع.
4. **معامل التحديد** R^2: يقيس مدى قرب البيانات من الخط المناسب، ويعطي فكرة عن نسبة التباين في المتغير التابع التي يمكن تفسيرها بواسطة المتغيرات المستقلة في النموذج.

الشبكات العصبية قادرة على تعلم التمثيلات المعقدة للبيانات، مما يجعلها أداة قوية لحل المشاكل

2.3.1 كيف يمكن للشبكات العصبية تقديم تقديرات للانحدار الخطي

الشبكات العصبية قادرة على تعلم تمثيلات الميزات المعقدة من البيانات ويمكنها تقديم تقديرات فعالة للعلاقات الخطية بين المتغيرات [?]. يمكن استخدام شبكة عصبية ذات طبقة واحدة أو أكثر لهذا الغرض [17].

- الشبكة العصبية ذات الطبقة الواحدة تُعد الأبسط لتمثيل العلاقات الخطية، حيث تتكون من طبقة واحدة من الخلايا العصبية تتطابق مع المتغيرات المستقلة.
- شبكات عصبية متعددة الطبقات يمكن أن تتعامل مع العلاقات غير الخطية المعقدة، ولكن يمكن تصميمها لتقديم تقديرات خطية بضبط الطبقة الخارجية لتكون خطية [11].

3.3.1 مقارنة بين النماذج الخطية التقليدية ونماذج الشبكات العصبية

- **قدرة التعميم:** الشبكات العصبية تتفوق في التعامل مع العلاقات المعقدة وغير الخطية، بينما النماذج الخطية فعالة في الحالات التي تكون العلاقة بين المتغيرات خطية بشكل واضح [13].
- **التعقيد والتفسيرية:** الشبكات العصبية غالبًا ما تكون أقل تفسيرية بسبب تعقيدها، على عكس النماذج الخطية التي تتميز ببساطتها وسهولة التفسير [?].
- **المرونة والتكيف:** الشبكات العصبية يمكنها التكيف تلقائيًا مع أنواع مختلفة من البيانات واستخراج الميزات المعقدة، بينما النماذج الخطية لديها قيود في هذا الجانب [17].
- **التحديات في التدريب:** الشبكات العصبية تتطلب كميات كبيرة من البيانات وموارد حسابية للتدريب، على عكس النماذج الخطية التي تكون أسهل في التدريب [11].

باب 2

تقدير الانحدار الخطي باستخدام الشبكات العصبية

1.2 مدخل

في سياق الانحدار الخطي باستخدام الشبكات العصبية، المقدر ليس معرفا بالمعنى التقليدي كما في النماذج الإحصائية الكلاسيكية. بدلاً من ذلك، المقدر هو النموذج الناتج عن تدريب الشبكة العصبية نفسها. الشبكات العصبية تتعلم تمثيل العلاقة بين المتغيرات المستقلة (المدخلات) والمتغير التابع (المخرجات) من خلال ضبط وزن الروابط بين العقد في مختلف طبقات الشبكة خلال عملية التدريب.

عند استخدام الشبكات العصبية لتقدير الانحدار الخطي، الهدف هو تدريب الشبكة بحيث تتوصل إلى تقديرات لوزن الروابط (المعادلة لمعاملات الانحدار في النموذج الخطي) التي تمكن الشبكة من تقديم توقعات دقيقة للمتغير التابع بناءً على المتغيرات المستقلة.

- طبقة المدخل: تتألف من عدة عقد، كل منها يمثل متغيراً مستقلاً.
- طبقة الخروج: عقدة واحدة تمثل المتغير التابع (القيمة المتوقعة).
- الوزن: الروابط بين طبقة المدخل وطبقة الخروج تحمل وزناً، وهذه الأوزان تُعدل خلال عملية التدريب.

عملية التدريب تشمل استخدام خوارزمية مثل الانتشار العكسي وتقنيات تحسين مثل مُحسّن آدم أو الانحدار التدريجي العشوائي لتحديث الأوزان بشكل يقلل من الخطأ بين القيم المتوقعة من الشبكة والقيم الفعلية في بيانات التدريب.

بعد التدريب، يُستخدم النموذج لتقدير قيمة المتغير التابع استنادًا إلى قيم جديدة للمتغيرات المستقلة، حيث يعمل النموذج كمقدر للعلاقة بين هذه المتغيرات.

2.2 مقدمة

الانحدار الخطي هو طريقة إحصائية تستخدم لفهم العلاقة بين متغير تابع واحد (عادة ما يُسمى y) ومتغير (أو متغيرات) مستقلة (المُسمى x). الهدف هو تقدير القيم المستقبلية للمتغير التابع استنادًا إلى قيم المتغيرات المستقلة. في حين أن الانحدار الخطي التقليدي يقدم نموذجًا بسيطًا ومباشرًا، توفر الشبكات العصبية أدوات قوية لتقدير الانحدار الخطي مع التعامل مع العلاقات المعقدة وغير الخطية بين المتغيرات.

3.2 الجانب النظري

الانحدار الخطي المتعدد

تعبر المعادلة العامة للانحدار الخطي المتعدد عن العلاقة بين مجموعة من المتغيرات المستقلة ومتغير التابع. الصيغة الرياضية للانحدار الخطي المتعدد هي كالتالي:

حالة مع عنصر الثابت المنفصل

$$Y = XW + B$$

حيث:

- Y هي مصفوفة النواتج بأبعاد $m \times 1$.
- X هي مصفوفة المتغيرات المستقلة بأبعاد $m \times n$.
- W هي مصفوفة الأوزان بأبعاد $n \times 1$.
- B هي مصفوفة الثوابت بأبعاد $m \times 1$.

حالة مع عنصر الثابت المدمج

$$Y = XW$$

في هذه الحالة، X تشمل عمودًا إضافيًا من الواحدات لتمثيل الثوابت و W يشمل وزن الانحياز كأول عنصر.

4.2 دوال التنشيط في الشبكات العصبية

دوال التنشيط هي مكونات أساسية في الشبكات العصبية الاصطناعية. تُستخدم لتحويل الإشارات أو المدخلات التي تتلقاها الخلايا العصبية إلى مخرجات. الوظيفة الرئيسية لدوال التنشيط هي إدخال اللاخطية إلى النموذج، مما يمكّن الشبكات العصبية من تعلم الأنماط المعقدة. [12]

أنواع دوال التنشيط

- **الدالة الخطية** : [8]

$$f(x) = x$$

تعتبر الدالة الخطية من أبسط أشكال دوال التنشيط، حيث أنها تمرر الإشارة دون تغيير. تستخدم عادة في الطبقة الأخيرة لمهام الانحدار حيث نحتاج إلى مخرج مستمر وغير محدود.

- **دالة الخطوة** : [8]

$$f(x) = \begin{cases} 0 & if \quad x < 0 \\ 1 & if \quad x \geq 0 \end{cases}$$

دالة الخطوة تحول كل القيم السالبة إلى صفر والموجبة أو المساوية للصفر إلى واحد. تُستخدم في نماذج الشبكات العصبية الثنائية، خاصة في المصنفات الثنائية.

- **دالة سيغمويد** : [8]

$$f(x) = \frac{1}{1 + e^{-x}}$$

دالة سيغمويد تحول القيم المدخلة إلى مدى بين 0 و1، مما يجعلها خياراً ممتازاً لنماذج التصنيف الثنائي حيث تُستخدم في الطبقة الأخيرة للشبكة لتمثيل احتمالية.

- **دالة التانجنت الزائدي** : [10]

$$f(x) = \tanh(x) = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}}$$

تعمل دالة التانجنت الزائدي بشكل مشابه لدالة سيغمويد ولكنها تعيد قيمًا بين -1 و1، مما يجعلها أكثر فعالية في بعض الطبقات الخفية بسبب تمركزها حول الصفر.

- **دالة التصحيح الخطي ريلو**: [12]

$$f(x) = \max(0, x)$$

تعد ReLU واحدة من أكثر دوال التنشيط شيوعًا في الشبكات العصبية العميقة، حيث تحل مشكلة الانفجار والاختفاء التدريجي للتدرجات عبر إدخال اللاخطية دون التأثير على القيم الإيجابية.

- **دالة ليكو-ريلو** : [10]

$$f(x) = \begin{cases} x & if \quad x \geq 0 \\ \alpha x & if \quad x < 0 \end{cases}$$

تعتبر تعديل لدالة ReLU حيث تسمح بتدفق صغير من التدرج حتى عندما يكون المدخل سالبًا، مما يساعد في تجنب مشكلة موت النيورونات التي قد تحدث في ReLU.

- **دالة سويش** : [12]

$$f(x) = x \cdot \sigma(x) = \frac{x}{1 + e^{-x}}$$

دالة سويش هي دالة تنشيط غير خطية تعمل بشكل جيد في مجموعة متنوعة من الشبكات العصبية. يُعتقد أنها تتفوق على ReLU في بعض السياقات بسبب خصائصها التي تجمع بين خصائص دالة سيغمويد و ReLU.

اشتقاقات دوال التنشيط

- **الدالة الخطية**

$$f(x) = x$$

اشتقاقها:

$$f'(x) = 1$$

الشكل التفاضلي:

$$\frac{df}{dx} = 1$$

- **دالة الخطوة**

$$f(x) = \begin{cases} 0 & if \quad x < 0 \\ 1 & if \quad x \geq 0 \end{cases}$$

اشتقاقها:

$f'(x) = 0$ لكل $x \neq 0$

وتكون دالة الدلتا ديراك $\delta(x)$ عند $x = 0$. الشكل التفاضلي:

$$\frac{df}{dx} = \delta(x)$$

- **دالة سيغمويد**

$$f(x) = \frac{1}{1 + e^{-x}}$$

اشتقاقها:

$$f'(x) = \frac{e^{-x}}{(1 + e^{-x})^2} = f(x)(1 - f(x))$$

الشكل التفاضلي:

$$\frac{df}{dx} = f(x)(1 - f(x))$$

- **دالة التانجنت الزائدي**

$$f(x) = \tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

اشتقاقها:

$$f'(x) = 1 - \tanh^2(x) = \text{sech}^2(x)$$

الشكل التفاضلي:

$$\frac{df}{dx} = 1 - \tanh^2(x)$$

- **دالة التصحيح الخطي ريلو**

$$f(x) = \max(0, x)$$

اشتقاقها:

$$f'(x) = \begin{cases} 1 & if \quad x > 0 \\ 0 & if \quad x < 0 \end{cases}$$

الشكل التفاضلي:

$$\frac{df}{dx} = \begin{cases} 1 & if \quad x > 0 \\ 0 & if \quad x < 0 \end{cases}$$

- **دالة ليكو-ريلو**

$$f(x) = \begin{cases} x & if \quad x \geq 0 \\ \alpha x & if \quad x < 0 \end{cases}$$

اشتقاقها:

$$f'(x) = \begin{cases} 1 & if \quad x \geq 0 \\ \alpha & if \quad x < 0 \end{cases}$$

الشكل التفاضلي:

$$\frac{df}{dx} = \begin{cases} 1 & if \quad x \geq 0 \\ \alpha & if \quad x < 0 \end{cases}$$

- **دالة سويش**

$$f(x) = x \cdot \sigma(x) = \frac{x}{1 + e^{-x}}$$

اشتقاقها باستخدام قاعدة الضرب:

$$f'(x) = \sigma(x) + x\sigma'(x) = \sigma(x) + x\sigma(x)(1 - \sigma(x))$$

الشكل التفاضلي:

$$\frac{df}{dx} = \sigma(x) + x\sigma(x)(1 - \sigma(x))$$

اختيار دالة التنشيط

اختيار دالة التنشيط (Activation-Function) هو خطوة مهمة وحاسمة في تصميم الشبكات العصبية. يلعب الاختيار المناسب لدالة التنشيط دوراً كبيراً في أداء الشبكة العصبية وكفاءتها. فيما يلي بعض العوامل التي يجب مراعاتها عند اختيار دالة التنشيط[19]:

- **قابلية التفاضل :(Dffierentiability)** يجب أن تكون دالة التنشيط قابلة للتفاضل لتتمكن من استخدام خوارزميات التدريب القائمة على تدرج المشتقات (مثل خوارزمية الانحدار العكسي).

- **التشبع:** بعض دوال التنشيط تعاني من مشكلة التشبع، حيث يتسبب التشبع في توقف القيم المخرجة عند حد معين، مما يؤدي إلى توقف تعلم النموذج. من الأمثلة على ذلك دالتي الsigmoid وtanh. في هذه الحالة، تعتبر دالة ReLU خياراً جيداً لأنها لا تعاني من هذه المشكلة.

- **الميل للإنفجار والتضاؤل :** تعتبر مشكلة تلاشي التدرج واحدة من المشاكل الشائعة مع الدوال مثل sigmoid وtanh، حيث يمكن أن يصبح التدرج صغيراً جداً مما يؤدي إلى بطء تعلم النموذج. دالة ReLU تساعد في تخفيف هذه المشكلة.

- **التقارب :** دوال التنشيط المختلفة يمكن أن تؤثر على سرعة التقارب في عملية التدريب. على سبيل المثال، دالة ReLU تساعد في سرعة التقارب ولكن يمكن أن تسبب مشكلة "الموت" (Dead-Neurons) حيث تتوقف بعض الوحدات العصبية عن الاستجابة.

- **الانتشار العكسي :(Backpropagation)** يجب اختيار دالة تنشيط تسهل عملية الانتشار العكسي للتدرج. دوال مثل ReLU وLeaky-ReLU وELUتساهم في تحقيق ذلك بكفاءة.

- **الدقة :** في بعض الأحيان، يمكن أن تؤثر دالة التنشيط على دقة النموذج. لذا يجب اختبار عدة دوال تنشيط لتحديد أي منها يحقق أفضل أداء للنموذج المعني.

5.2 التدريب والانتشار الأمامي والعكسي

1.5.2 الانتشار الأمامي

في الانتشار الأمامي، نحسب القيمة المتوقعة باستخدام الأوزان الحالية. النموذج العام للشبكة العصبية، مع استخدام دالة التنشيط ReLU في الطبقات الوسطى، يمكن أن يُعبر عنه كالتالي:

- **للطبقة الأولى:**

$$Z^{[1]} = W^{[1]}X + b^{[1]}$$
$$A^{[1]} = \max(0, Z^{[1]})$$

- **لكل طبقة وسطى** $l = 2, \ldots, L-1$:

$$Z^{[l]} = W^{[l]}A^{[l-1]} + b^{[l]}$$
$$A^{[l]} = \max(0, Z^{[l]})$$

- **للطبقة الأخيرة:**

$$Z^{[L]} = W^{[L]}A^{[L-1]} + b^{[L]}$$
$$\hat{Y} = Z^{[L]}$$

التفسير

- **الطبقة الأولى:** تقوم هذه الطبقة بحساب مجموع الإدخالات الموزونة والتحيز، وتطبيق دالة ReLU. دالة ReLU تسمح بمرور القيم الإيجابية دون تغيير وتحول القيم السلبية إلى صفر.

- **الطبقات الوسطى:** تستمر كل طبقة في معالجة البيانات من الطبقة السابقة بنفس الطريقة، مما يزيد من تعقيد التمثيلات التي يمكن للشبكة التقاطها.

- **الطبقة الأخيرة:** في الطبقة الأخيرة، غالبًا ما تستخدم دالة الهوية للتنبؤ بقيم مستمرة أو دالة سيغمويد/softmax للتصنيف. هذه الطبقة تنتج القيم المتوقعة النهائية للشبكة.

2.5.2 تهيئة He [15]

تهيئة He هي تقنية لتهيئة الأوزان في الشبكات العصبية، تم تطويرها بواسطة Kaiming-He وزملائه بهدف تحسين استقرار وسرعة تدريب الشبكات العصبية، خاصة عند استخدام دالة التنشيط ReLU . الهدف من هذه التهيئة هو الحفاظ على التباين عبر كل طبقة في الشبكة لتجنب مشاكل الانفجار أو الاختفاء التدريجي للتدرجات أثناء التدريب. المعادلة الأساسية لتهيئة He هي:

$$W \sim \mathcal{N}\left(0, \sqrt{\frac{2}{n}}\right)$$

حيث:

- W هي الأوزان المبدئية.
- $\mathcal{N}(0, \sigma^2)$ يمثل التوزيع الطبيعي (الغاوسي) بمتوسط 0 وانحراف معياري σ.
- n هو عدد الوحدات الداخلة أو عدد المدخلات للعصبون.

حيث g هي دالة التنشيط،

الأساس الرياضي لتهيئة He

تهيئة He تعتمد على استخدام توزيع طبيعي لتهيئة الأوزان W بمتوسط صفر وتباين يعتمد على عدد المدخلات n:

$$W \sim \mathcal{N}(0, \sigma^2)$$

حيث:

$$\sigma = \sqrt{\frac{2}{n}}$$

لنفترض أن الإشارات المدخلة x لديها تباين σ_x^2. الإخراج من العصبون قبل تطبيق دالة التنشيط هو $z = \sum_{i=1}^{n} w_i x_i$. بناءً على استقلالية وتوزيع الأوزان والإدخالات:

$$\text{Var}(z) = \text{Var}\left(\sum_{i=1}^{n} w_i x_i\right) = \sum_{i=1}^{n} \text{Var}(w_i x_i)$$

$$\text{Var}(w_i x_i) = \text{Var}(w_i)\text{Var}(x_i) = \frac{2}{n}\sigma_x^2$$

$$\text{Var}(w_i x_i) = w_i^2 \text{Var}(x_i)$$

بافتراض أن $\text{Var}(x_i) = \sigma_x^2$ و $w_i = \sqrt{\frac{2}{n}}$:

$$\text{Var}(w_i x_i) = \left(\sqrt{\frac{2}{n}}\right)^2 \sigma_x^2 = \frac{2}{n}\sigma_x^2$$

المجموع الكلي للتباين يكون:

$$\text{Var}(z) = n\frac{2}{n}\sigma_x^2 = 2\sigma_x^2$$

هذا يعني أن تباين الإخراج z يُضاعف بمقدار 2 مقارنة بتباين الإدخال، مما يساعد في الحفاظ على استقرار التدرجات خلال التدريب، خصوصًا عند استخدام دالة التنشيط ريلو.

فوائد تهيئة He

- **تسريع التدريب:** تساعد هذه الطريقة في تسريع عملية التدريب عن طريق الحفاظ على تدفق الإشارات بشكل صحيح عبر الشبكة.
- **منع مشاكل التلاشي والانفجار:** تقلل تهيئة He من مشاكل تلاشي وانفجار القيم أثناء انتقال الإشارات عبر الطبقات.

3.5.2 الانتشار العكسي

لتحديث الأوزان والانحيازات، نستخدم الانتشار العكسي لحساب المشتقات الجزئية لدالة الخسارة بالنسبة لكل معامل:
دالة الخسارة يتم استخدام دالة الخسارة لقياس مدى جودة النموذج في التنبؤ بالقيم الحقيقية.

- **متوسط الخطأ التربيعي**

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - A^{[L](i)})^2 \tag{1}$$

- **الخطأ المطلق المتوسط**

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - A^{[L](i)})| \tag{2}$$

أهمية دالة الخسارة
دالة الخسارة توجه النموذج للتعلم بتحديد كيفية تعديل الوزن في الشبكة لتحقيق أداء أفضل، وتُستخدم لتقييم أداء النموذج خلال التدريب وبعده.

خطوات اشتقاق دالة التكلفة نبدأ بتعريف دالة التكلفة:

$$J(W,b) = \frac{1}{n}\sum_{i=1}^{n}(y^{(i)} - A^{[L](i)})^2$$

حيث:

- n هو عدد العينات.
- $y^{(i)}$ هو القيمة المستهدفة للعينة i.
- $A^{[L](i)}$ هو مخرجات الطبقة الأخيرة (L) للعينة i.

خطأ الطبقة الأخيرة خطأ الطبقة الأخيرة (L):

$$E^{[L](i)} = A^{[L](i)} - y^{(i)}$$

حيث:

- $E^{[L](i)}$ هو خطأ العينة i في الطبقة L.

اشتقاق دالة التكلفة بالنسبة للأوزان $W^{[L]}$ لاشتقاق دالة التكلفة بالنسبة للأوزان، نبدأ بالتعبير عن دالة التكلفة على النحو التالي:

$$\frac{\partial J}{\partial W^{[L]}} = \frac{2}{n}\sum_{i=1}^{n} E^{[L](i)}\frac{\partial E^{[L](i)}}{\partial W^{[L]}}$$

اشتقاق الخطأ $E^{[L](i)}$ **بالنسبة للأوزان** $W^{[L]}$ نبدأ بالتعبير عن $E^{[L](i)}$ في شكل مبسط:

$$E^{[L](i)} = A^{[L](i)} - y^{(i)}$$

وبما أن دالة التنشيط هي دالة الهوية:

$$A^{[L](i)} = Z^{[L](i)}$$

حيث:

$$Z^{[L](i)} = W^{[L]}A^{[L-1](i)} + b^{[L]}$$

إذًا،

$$E^{[L](i)} = W^{[L]}A^{[L-1](i)} + b^{[L]} - y^{(i)}$$

لاشتقاق $E^{[L](i)}$ بالنسبة للأوزان $W^{[L]}$:

$$\frac{\partial E^{[L](i)}}{\partial W^{[L]}} = \frac{\partial}{\partial W^{[L]}}(W^{[L]}A^{[L-1](i)} + b^{[L]} - y^{(i)})$$

بما أن $b^{[L]}$ و $y^{(i)}$ لايتعلقان ب $W^{[L]}$، فإن اشتقاقهما يساوي صفر:

$$\frac{\partial E^{[L](i)}}{\partial W^{[L]}} = \frac{\partial}{\partial W^{[L]}}(W^{[L]}A^{[L-1](i)})$$

وبما أن:

$$\frac{\partial}{\partial W^{[L]}}(W^{[L]}A^{[L-1](i)}) = A^{[L-1](i)}$$

إذًا،

$$\frac{\partial E^{[L](i)}}{\partial W^{[L]}} = A^{[L-1](i)}$$

تعويض المشتقات في اشتقاق دالة التكلفة الآن، يمكننا تعويض المشتقات في التعبير الأصلي لاشتقاق دالة التكلفة:

$$\frac{\partial J}{\partial W^{[L]}} = \frac{2}{n}\sum_{i=1}^{n} E^{[L](i)}A^{[L-1](i)}$$

اشتقاق دالة التكلفة بالنسبة للإزاحة $b^{[L]}$ نظرًا لأن الإزاحة $b^{[L]}$ لا تتأثر بتغيير دالة التنشيط، فإن اشتقاق الخطأ $E^{[L](i)}$ بالنسبة لـ $b^{[L]}$ هو:

$$\frac{\partial E^{[L](i)}}{\partial b^{[L]}} = 1$$

وبالتالي، اشتقاق دالة التكلفة بالنسبة للإزاحة $b^{[L]}$:

$$\frac{\partial J}{\partial b^{[L]}} = \frac{2}{n}\sum_{i=1}^{n} E^{[L](i)}$$

خلاصة الاشتقاقات اشتقاق دالة التكلفة بالنسبة للأوزان $W^{[L]}$:

$$\frac{\partial J}{\partial W^{[L]}} = \frac{2}{n}\sum_{i=1}^{n} E^{[L](i)}A^{[L-1](i)}$$

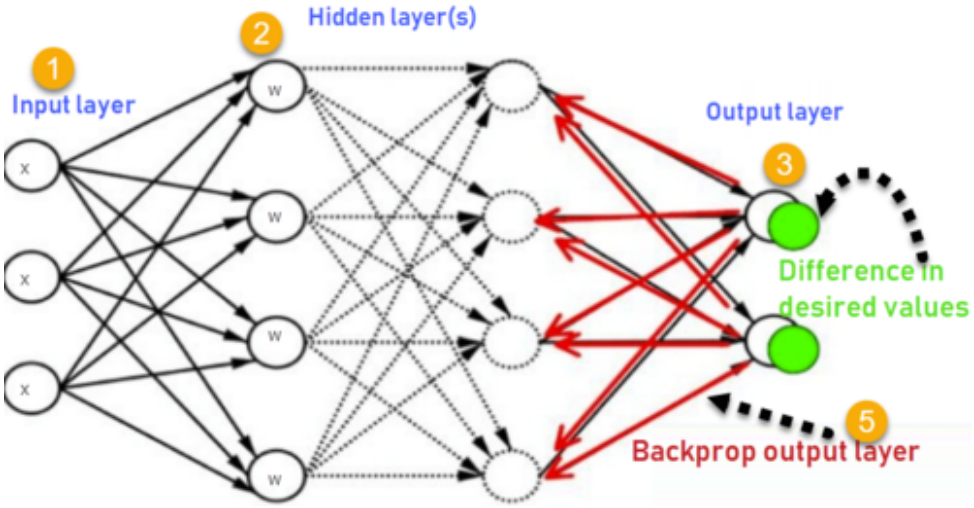
واشتقاق دالة التكلفة بالنسبة للإزاحة $b^{[L]}$:

$$\frac{\partial J}{\partial b^{[L]}} = \frac{2}{n}\sum_{i=1}^{n} E^{[L](i)}$$

بهذا الشكل، نكون قد حللنا الاشتقاقات اللازمة لتحديث الأوزان والتحيزات في الشبكة العصبية باستخدام خوارزمية الإنتشار العكسي.

خلاصة

الانتشار العكسي هو أساس تعلم الشبكات العصبية، حيث يتم استخدامه لتحديث الأوزان بطريقة تقلل من الخسارة، مما يساهم في تحسين قدرة النموذج على التنبؤ بدقة أعلى.



شكل 1.2: الإنتشار العكسي

6.2 تحسين النموذج [20]

تحسين النموذج هو عملية تعديل معلمات النموذج الرياضي أو الخوارزمية بحيث يتم تقليل دالة الخسارة وتحسين أداء النموذج على مجموعة معينة من البيانات. يهدف تحسين النموذج إلى العثور على القيم المثلى للمعلمات التي تؤدي إلى تحقيق أفضل أداء ممكن من حيث دقة التنبؤ أو تقليل الأخطاء. تُستخدم تقنيات التحسين على نطاق واسع في التعلم الآلي، والإحصاءات، والبحث العلمي لتحسين نتائج النماذج والتأكد من أنها تقدم أفضل أداء ممكن عند التنبؤ بالبيانات الجديدة.

يُمكن تحسين النموذج من خلال عدة تقنيات منها:

خوارزمية التحسين بالزخم

خوارزمية التحسين بالزخم هي تقنية تُستخدم في تحسين النماذج في مجال التعلم الآلي. تُضاف هذه التقنية إلى خوارزميات الانحدار التقليدية مثل انحدار التدرج البسيط لتسريع عملية التحسين وتحسين استقرارها. تعمل خوارزمية الزخم على تحديث معلمات النموذج بحيث تأخذ في الحسبان ليس فقط القيمة الحالية لتدرج دالة الخسارة، ولكن أيضًا الاتجاه السابق للتدرج. يُمثل التحديث في خوارزمية الزخم بالعلاقة التالية:

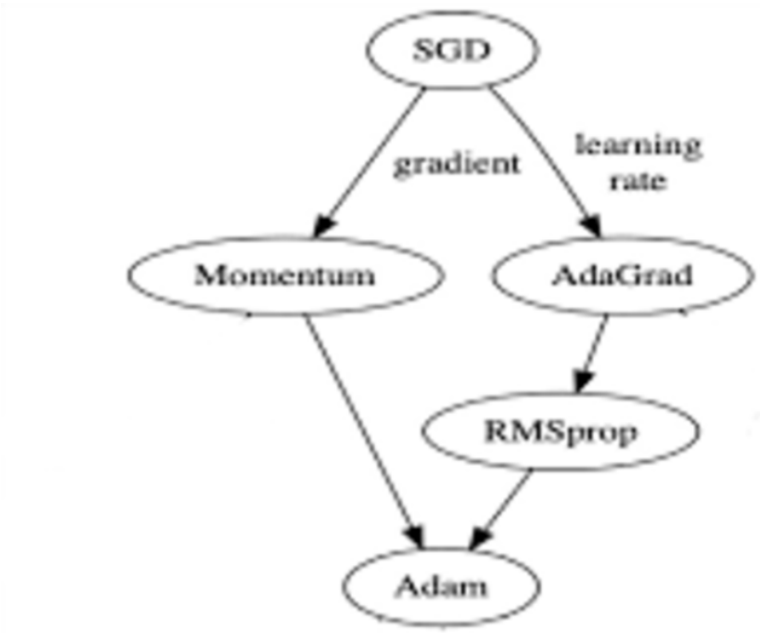
معادلة التحديث:

$$v_t = \gamma v_{t-1} + \eta \nabla_\theta J(\theta)$$

$$\theta = \theta - v_t$$

خوارزمية التحسين Adam

خوارزمية Adam هي واحدة من أكثر خوارزميات التحسين شيوعًا لتدريب شبكات العصبية. إنها تجمع بين أفكار من خوارزميتي RMSprop وSGD مع الزخم. تم تطويرها بواسطة Kingma و Ba في عام 2014، ويعتبر اسم "Adam`` اختصارًا لعبارة "Adaptive-Moments`` أو ``اللحظات التكيفية". تتميز هذه الخوارزمية بقدرتها على تعديل معدلات التعلم بشكل مستقل لكل معلمة من معلمات النموذج استنادًا إلى التقديرات التكيفية للحظات من الدرجة الأولى والثانية للتدرج.



شكل 2.2: خوارزمية التحسين Adam

مزايا خوارزمية آدم

- **تكيفية معدل التعلم:** تقوم آدم بتعديل معدلات التعلم لكل معلمة بناءً على تدرجات اللحظات التكيفية، مما يجعلها قادرة على التعامل مع المشاكل التي تنطوي على بيانات غير متجانسة أو خصائص غير ثابتة.
- **استقرار وتوازن في التدرج:** استخدام اللحظات التكيفية يقلل من تقلبات التدرج، مما يؤدي إلى تحسينات في استقرار وتوازن التحديثات.
- **سهولة التنفيذ والفعالية:** تعد آدم واحدة من الخوارزميات الأكثر استخدامًا في مجال تعلم الآلة نظرًا لسهولة تنفيذها وفعاليتها في تحسين نماذج الشبكات العصبية.

خطوات خوارزمية آدم

في بداية الأمر، نعرف تحديثات اللحظة الأولى والثانية كالتالي:

$$
\begin{aligned}
m_{(\theta)}^{[t]} &= \beta_1 m_{(\theta)}^{[t-1]} + (1-\beta_1)(\nabla_\theta J(\theta)),\\
m_W^{[t]} &= \beta_1 m_W^{[t-1]} + (1-\beta_1)\left(\frac{\partial J}{\partial W^{[L]}}\right),\\
m_b^{[t]} &= \beta_1 m_b^{[t-1]} + (1-\beta_1)\left(\frac{\partial J}{\partial b^{[L]}}\right),\\
v_{(\theta)}^{[t]} &= \beta_2 v_{(\theta)}^{[t-1]} + (1-\beta_2)(\nabla_\theta J(\theta))^2,\\
v_{(\theta)}^{[t]} &= \beta_2 v_{(\theta)}^{[t-1]} + (1-\beta_2)((\nabla_\theta J(\theta)) \odot (\nabla_\theta J(\theta))),\\
v_W^{[t]} &= \beta_2 v_W^{[t-1]} + (1-\beta_2)\left(\frac{\partial J}{\partial W^{[L]}} \odot \frac{\partial J}{\partial W^{[L]}}\right),\\
v_b^{[t]} &= \beta_2 v_b^{[t-1]} + (1-\beta_2)\left(\frac{\partial J}{\partial b^{[L]}} \odot \frac{\partial J}{\partial b^{[L]}}\right).
\end{aligned}
$$

حيث:

- $\nabla_\theta J(\theta)$ هو التدرج في الخطوة الزمنية t.
- $m_{(\theta)}^{[t]}$ هو تقدير اللحظة الأولى (أي متوسط الحركة).
- $v_{(\theta)}^{[t]}$ هو تقدير اللحظة الثانية (أي متوسط الحركة المربعة).
- β_1 و β_2 هما معاملات التحكم في تلاشي اللحظات,و$\odot$ يشير إلى الضرب التكويني للمصفوفات (عنصر بعنصر)

لتصحيح التحيز في التقديرات الأولية، التي تكون متحيزة نحو الصفر بسبب التهيئة، نستخدم:

$$
\begin{aligned}
\hat{m}_{(\theta)}^{[t]} &= \frac{m_{(\theta)}^{[t]}}{1-\beta_1^t},\\
\hat{v}_{(\theta)}^{[t]} &= \frac{v_{(\theta)}^{[t]}}{1-\beta_2^t}.
\end{aligned}
$$

حيث:

- $\hat{m}_{(\theta)}^{[t]}$ و $\hat{v}_{(\theta)}^{[t]}$ هي تصحيحات التحيز لـ $m_{(\theta)}^{[t]}$ و $v_{(\theta)}^{[t]}$ على التوالي.
- β_1 و β_2 هي معاملات تقدير الزخم.
- η هو معدل التعلم.
- ϵ هو مصطلح التحسين لتجنب القسمة على صفر.

معادلات تصحيح الوزن والتحيز:

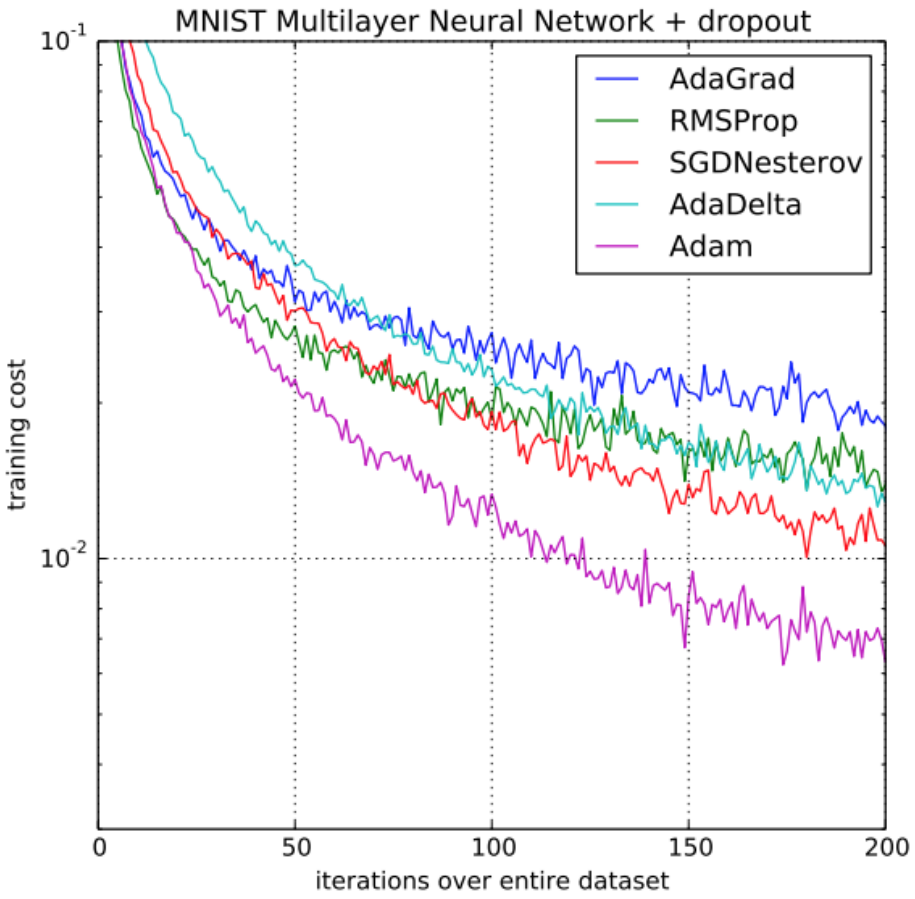
$$\theta^{[l]} = \theta^{[l]} - \alpha \frac{\hat{m}_{(\theta)}^{[l]}}{\sqrt{\hat{v}_{(\theta)}^{[l]}} + \epsilon},$$
$$W^{[l]} = W^{[l]} - \alpha \frac{\hat{m}_{W}^{[l]}}{\sqrt{\hat{v}_{W}^{[l]}} + \epsilon},$$
$$b^{[l]} = b^{[l]} - \alpha \frac{\hat{m}_{b}^{[l]}}{\sqrt{\hat{v}_{b}^{[l]}} + \epsilon}.$$

برهان تصحيح التحيز في متوسط التدرجات

- $m_{(\theta)}^{[t]}$ يحسب كمتوسط مرجح للتدرجات، حيث يدير β_1 التوازن بين التدرج الحالي g_t والتدرجات السابقة.
- بما أن $m_{(\theta)}^{[t]}$ يبدأ من الصفر ويتم تحديثه بشكل تراكمي باستخدام عامل التضاؤل β_1, بعد t خطوات, $m_{(\theta)}^{[t]}$ سيكون مقللاً بمعامل $(1 - \beta_1^t)$.
- لذلك، لتصحيح التحيز في $m_{(\theta)}^{[t]}$, يتم تقسيم $m_{(\theta)}^{[t]}$ على $(1 - \beta_1^t)$ للحصول على التقدير المصحح $\hat{m}_{(\theta)}^{[t]}$.

خواص خوارزمية Adam:

- خوارزمية Adam تجمع بين فوائد مُعدِّل اللحظة الأولى ومُعدِّل اللحظة الثانية.
- تُعد فعّالة لأنها تُقلل من مشكلة اختيار معدل التعلم المناسب وتحسن الاستقرار في عملية التحسين، مما يجعلها مفيدة بشكل خاص عند تحديث الانحيازات.
- خوارزمية آدم هي تقنية لتحسين النزول الأسرع المستخدمة في تدريب الشبكات العصبية وغيرها من التطبيقات التحسينية.
- تستخدم آلية لتعديل معدل التعلم بشكل تكيفي لكل معامل، مما يساعد في تسريع عملية التقارب.



شكل 3.2: تقارب دالة الخطأ لبيانات التدريب

ملاحظة:

تصحيح التحيز ضروري لأن $m_{(\theta)}^{[t]}$ و $v_{(\theta)}^{[t]}$ يبدآن متحيزين نحو الصفر. تصحيح التحيز يقوم بتعديل هذه التقديرات لتكون أقرب إلى قيمها الحقيقية مع تقدم الزمن.

مع زيادة t, β_1^t و β_2^t تقتربان من الصفر, مما يجعل $1 - \beta_1^t$ و $1 - \beta_2^t$ تقتربان من 1. هذا يؤدي إلى تقليل تأثير التحيز في التقديرات مع الزمن.

تحديث الأوزان والانحيازات باستخدام خوارزمية آدم

الانتشار الأمامي في الشبكة العصبية:

- **للطبقة الأولى:**

$$Z^{[1]} = W^{[1]}X + b^{[1]}$$

$$A^{[1]} = \max(0, Z^{[1]}) \quad \text{(استخدام دالة ReLU كدالة تنشيط)}$$

- **لكل طبقة وسطى** $l = 2, \ldots, L-1$:

$$Z^{[l]} = W^{[l]}A^{[l-1]} + b^{[l]}$$

$$A^{[l]} = \max(0, Z^{[l]})$$

- **للطبقة الأخيرة:**

$$Z^{[L]} = W^{[L]} A^{[L-1]} + b^{[L]}$$

$$\hat{Y} = Z^{[L]} \quad \text{(استخدام دالة الهوية في الطبقة الأخيرة للحصول على الناتج)}$$

دالة الخسارة:

$$J(W, b) = \frac{1}{n} \sum_{i=1}^{n} (y^{(i)} - A^{[L](i)})^2$$

الخطأ في الطبقة الأخيرة:

$$E^{[L](i)} = A^{[L](i)} - y^{(i)}$$

اشتقاق دالة التكلفة بالنسبة للأوزان $W^{[L]}$ والانحيازات $b^{[L]}$:

$$\frac{\partial J}{\partial W^{[L]}} = \frac{2}{n} \sum_{i=1}^{n} E^{[L](i)} A^{[L-1](i)}$$

$$\frac{\partial J}{\partial b^{[L]}} = \frac{2}{n} \sum_{i=1}^{n} E^{[L](i)}$$

الخطوات المتتالية للطبقات المتوسطة

بالنسبة لكل طبقة $l = L-1, L-2, \ldots, 1$:

- **الانتشار الأمامي:**

$$Z^{[l]} = W^{[l]} A^{[l-1]} + b^{[l]}$$

$$A^{[l]} = \max(0, Z^{[l]})$$

- **حساب المشتقات:**

$$\frac{\partial J}{\partial Z^{[l]}} = \frac{\partial J}{\partial A^{[l]}} \odot g'(Z^{[l]})$$

$$\frac{\partial J}{\partial Z^{[l-1]}} = (W^{[l]T} \frac{\partial J}{\partial Z^{[l]}}) \odot g'(Z^{[l-1]})$$

- **تحديث الأوزان والانحيازات باستخدام خوارزمية آدم:**

$$m_W^{[l]} = \beta_1 m_W^{[l]} + (1 - \beta_1) \frac{\partial J}{\partial W^{[l]}}$$

$$v_W^{[l]} = \beta_2 v_W^{[l]} + (1 - \beta_2) \left(\frac{\partial J}{\partial W^{[l]}} \right)^2$$

$$\hat{m}_W^{[l]} = \frac{m_W^{[l]}}{1 - \beta_1^t}$$

$$\hat{v}_W^{[l]} = \frac{v_W^{[l]}}{1 - \beta_2^t}$$

$$W^{[l]} = W^{[l]} - \alpha \frac{\hat{m}_W^{[l]}}{\sqrt{\hat{v}_W^{[l]}} + \epsilon}$$

$$b^{[l]} = b^{[l]} - \alpha \frac{\hat{m}_b^{[l]}}{\sqrt{\hat{v}_b^{[l]}} + \epsilon}$$

الخطوات المتتالية للطبقة الأولى

- **الانتشار الأمامي:**

$$Z^{[1]} = W^{[1]}X + b^{[1]}$$
$$A^{[1]} = \max(0, Z^{[1]})$$

- **حساب المشتقات:**

$$\frac{\partial J}{\partial Z^{[1]}} = \frac{\partial J}{\partial A^{[1]}} \odot g'(Z^{[1]})$$

- **تحديث الأوزان والانحيازات باستخدام خوارزمية آدم:**

$$m_W^{[1]} = \beta_1 m_W^{[1]} + (1-\beta_1)\frac{\partial J}{\partial W^{[1]}}$$

$$v_W^{[1]} = \beta_2 v_W^{[1]} + (1-\beta_2)\left(\frac{\partial J}{\partial W^{[1]}}\right)^2$$

$$\hat{m}_W^{[1]} = \frac{m_W^{[1]}}{1-\beta_1^t}$$

$$\hat{v}_W^{[1]} = \frac{v_W^{[1]}}{1-\beta_2^t}$$

$$W^{[1]} = W^{[1]} - \alpha\frac{\hat{m}_W^{[1]}}{\sqrt{\hat{v}_W^{[1]}} + \epsilon}$$

$$b^{[1]} = b^{[1]} - \alpha\frac{\hat{m}_b^{[1]}}{\sqrt{\hat{v}_b^{[1]}} + \epsilon}$$

7.2 تقنيات منع الإفراط في التعلم [16],[12]

التنظيم:

التنظيم هو تقنية تُستخدم في نمذجة التعلم الآلي لتقليل الإفراط في التكيف (overfitting) وتحسين أداء النموذج على البيانات الجديدة. يهدف التنظيم إلى إضافة قيود أو تعديلات على النموذج بحيث لا يتعلم تفاصيل ضجيج البيانات بشكل مفرط. يتم تحقيق ذلك عادةً بإضافة مصطلح عقوبة إلى دالة الخسارة الأصلية. أكثر أنواع التنظيم شيوعًا هي تنظيم L1 وتنظيم L2.
تضيف عقوبة على مقدار الأوزان في النموذج، مما يشجع على الأوزان الأصغر وبالتالي نحصل على نموذج أبسط والذي يقاوم الإفراط في التعلم.

التنظيم :L1

مصطلح L1 يضيف المجموع المطلق للأوزان إلى دالة الخسارة لإنتاج نماذج أكثر نحافة. الشكل الرياضي:

$$J_{L1}(W,b) = J(W,b) + \lambda \sum_{i=1}^{n} |w_i|$$

حيث:

- $J(W,b)$ هي دالة الخسارة الأصلية.
- λ هو معامل التسوية.
- $\sum_{i=1}^{n} |w_i|$ هو مجموع المطلق لجميع الوزنيات.

التنظيم التربيعي :L2

مصطلح L2 يضيف مربع أوزان إلى دالة الخسارة لتقليل حجمها بشكل موحد. الشكل الرياضي:

$$J_{L2}(W,b) = J(W,b) + \lambda \sum_{i=1}^{n} w_i^2$$

حيث:

- $J(W,b)$ هي دالة الخسارة الأصلية.
- λ هو معامل التسوية.
- $\sum_{i=1}^{n} w_i^2$ هو مجموع مربعات الوزنيات.

التوقف المبكر :

هو تقنية تُستخدم أثناء تدريب النماذج التنبؤية، خاصة في مجال التعلم الآلي العميق، لمنع حدوث ظاهرة الإفراط في التكيف مع بيانات التدريب.

كيف يعمل التوقف المبكر

1. **الرصد أثناء التدريب:** خلال عملية التدريب، يتم مراقبة أداء النموذج ليس فقط على بيانات التدريب ولكن أيضاً على مجموعة من البيانات المنفصلة تُعرف ببيانات التحقق أو التحقق الصليبي.

2. **تحديد المعيار والصبر:** يُحدد معيار معين (مثل دقة النموذج أو الخطأ في التحقق) لتقييم متى يجب التوقف عن التدريب. كما يتم تحديد "الصبر"، وهو عدد الدورات التدريبية التي يجب أن ينتظرها المدرب قبل أن يتخذ قراراً بالتوقف إذا لم يلاحظ تحسناً.

3. **التوقف عند عدم التحسن:** إذا لم يتحسن أداء النموذج على بيانات التحقق لعدد الدورات المحدد بالصبر، فإن عملية التدريب تتوقف.

فوائد التوقف المبكر

- **تحسين التعميم:** التوقف المبكر يساعد النموذج على التعميم بشكل أفضل على بيانات جديدة، وهو ما يعزز فعاليته في الاستخدام العملي.
- **توفير الوقت والموارد:** من خلال إنهاء التدريب في وقت مبكر، يتم توفير الموارد والوقت الذي كان سيُستهلك في تدريبات غير ضرورية.

التحقق المتقاطع :

يقسم هذا الأسلوب البيانات إلى عدة أقسام (مثلاً، 5 أو 10)، ويستخدم كل قسم مرة واحدة كمجموعة اختبار بينما يستخدم الأقسام الأخرى للتدريب. يساعد ذلك في ضمان أن النموذج قادر على التعميم جيدًا عبر البيانات غير المرئية. يمكن استخدام التحقق المتقاطع لتقييم كيفية أداء النموذج على مجموعة بيانات غير مرئية لتقليل الإفراط في التعلم. لا يرتبط بمعادلة مباشرة ولكن بطريقة لتقسيم البيانات وتقييم النموذج.

زيادة البيانات :

زيادة حجم البيانات هي تقنية تُستخدم لتحسين أداء النماذج في التعلم الآلي عبر زيادة كمية البيانات المتاحة للتدريب من خلال إنشاء بيانات تدريب جديدة من البيانات الموجودة.

أساليب زيادة حجم البيانات

في مجال معالجة الصور، يمكن تطبيق العديد من التقنيات مثل:

- تدوير الصورة
- تغيير الحجم
- قص الصورة
- تغيير السطوع والتباين

في مجال معالجة اللغات الطبيعية، يمكن تطبيق تقنيات مثل:

- تعديل الترتيب الجمل
- استخدام المترادفات

فوائد زيادة حجم البيانات

- **تحسين الأداء:** تساعد زيادة حجم البيانات في تحسين أداء النماذج بتوفير بيانات تدريب أكثر تنوعاً.
- **منع الإفراط في التكيف:** تقلل هذه التقنية من خطر الإفراط في التكيف حيث يصبح النموذج قادراً على التعميم بشكل أفضل على بيانات جديدة.

تطبيق في تقدير الانحدار الخطي المتعدد

في سياق تقدير الانحدار الخطي المتعدد باستخدام الشبكات العصبية، يمكن استخدام زيادة البيانات لتحسين دقة النموذج وتقليل التحيز. على سبيل المثال، يمكن إنشاء نسخ جديدة من البيانات الأصلية مع بعض التعديلات الطفيفة لضمان تنوع البيانات وزيادة حجم مجموعة التدريب. يمكن استخدام تقنيات مثل إضافة ضوضاء صغيرة للبيانات أو إنشاء مجموعات فرعية عشوائية من البيانات الأصلية. هذه الأساليب تساعد في تدريب نموذج أكثر فعالية وقادر على التعميم بشكل أفضل على بيانات جديدة.

تقليم الشبكة :

تقليم الشبكة هو عملية تهدف إلى تقليل حجم وتعقيد الشبكة العصبية عن طريق إزالة الأوزان والوصلات غير الضرورية. هذا يساعد في تحسين كفاءة النموذج دون التضحية الكبيرة بالأداء.

أساليب تقليم الشبكة

هناك عدة طرق لتقليم الشبكة العصبية، منها:

- **التقليم المستند إلى الأوزان:** إزالة الأوزان التي تكون قيمتها أقل من عتبة معينة.
- **التقليم المستند إلى الوحدات:** إزالة الوحدات العصبية (neuron) بأكملها إذا كانت مساهمتها في الناتج الإجمالي للشبكة ضئيلة.
- **التقليم المستند إلى التكرار:** إزالة الوصلات المكررة أو المتشابهة.

فوائد تقليم الشبكة

- **تحسين الكفاءة الحسابية:** تقليل عدد العمليات الحسابية المطلوبة، مما يزيد من سرعة النموذج.
- **تقليل استخدام الذاكرة:** خفض حجم النموذج، مما يجعل تخزينه ونقله أكثر كفاءة.
- **منع الإفراط في التكيف:** يمكن أن يساعد تقليم الشبكة في تجنب الإفراط في التكيف عن طريق إزالة العناصر الزائدة.

تطبيق تقليم الشبكة في تقدير الانحدار الخطي المتعدد

في سياق تقدير الانحدار الخطي المتعدد باستخدام الشبكات العصبية، يمكن لتقليم الشبكة أن يساعد في تحسين أداء النموذج بشكل كبير. عند تقليم الشبكة، يمكن تحديد الأوزان والوصلات الأقل أهمية وإزالتها، مما يساعد في تقليل التعقيد الكلي للنموذج وتحسين كفاءته.

8.2 تقييم النموذج

- **متوسط الخطأ التربيعي (MSE):**

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - A^{[L](i)})^2 \tag{3}$$

- **الخطأ المطلق المتوسط (MAE):**

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - A^{[L](i)})| \tag{4}$$

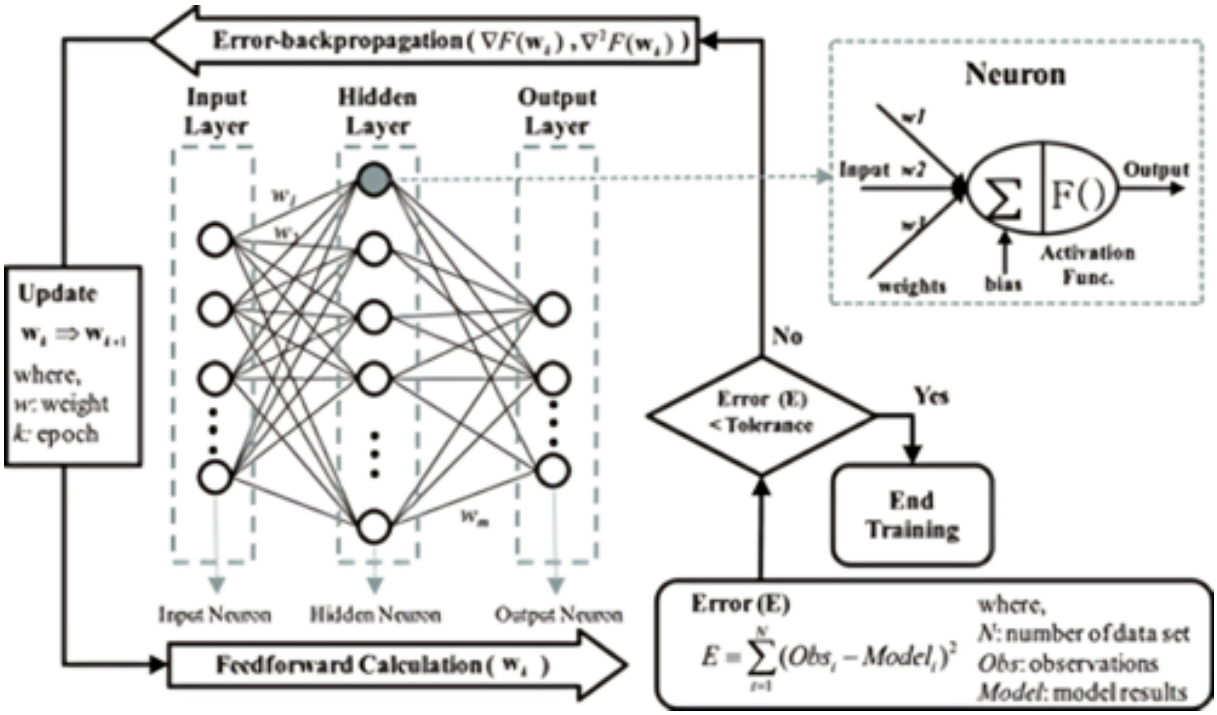
- **الجذر التربيعي للخطأ التربيعي الوسطي (RMSE):**

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - A^{[L](i)})^2} \tag{5}$$

- **معامل الارتباط R^2:**

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - A^{[L](i)})^2}{\sum_{i=1}^{n}(y_i - \overline{y})^2} \tag{6}$$

بعد تدريب النموذج، نقيم أدائه باستخدام مجموعة الاختبار ونحسب مقاييس مثل ،MSE ،RMSE و R^2 لتحديد مدى دقة النموذج.



شكل 4.2: شرح طريقة عمل الشبكة العصبية

باب 3

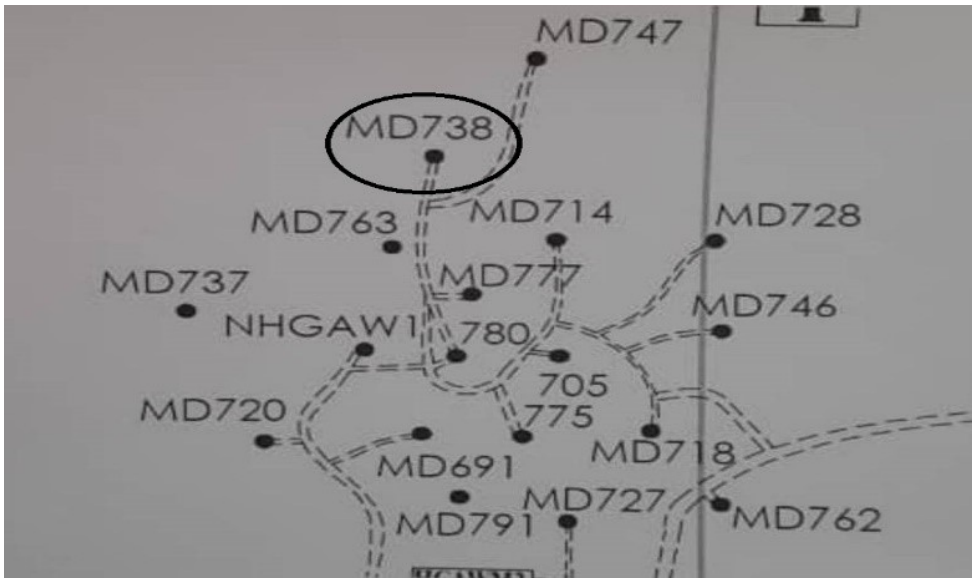
برمجة شبكة عصبية لتقدير الانحذار الخطي المتعدد بإستخدام لغة البرمجة R

مقدمة

النفاذية هي معلمة مهمة لتقييم قدرة الخزان على إنتاج الهيدروكربونات، لأنها تحدد السرعة التي يمكن للسوائل أن تتحرك بها عبر الصخر. قد تجعل النفاذية المنخفضة استخراج الهيدروكربونات صعباً وتستلزم تقنيات تحفيز مثل التكسير الهيدروليكي لزيادة الإنتاج. هذا الفصل يحتوي على مجموعة من النتائج التي تم الحصول عليها بواسطة طريقة التنبؤ بالنفاذية .

تحليل البيانات

البئر العمودي الموجود في MD738 يقع في منطقة حاسي مسعود - منطقة HZS المعقدة (خارج المنطقة الجنوبية). هذا البئر يأتي في إطار تطوير حقل حاسي مسعود الجنوبي الغربي، بهدف استهداف الخزان كومبو-الأوردوفيشي. استخدمنا مجموعة بيانات من القلب وتسجيل النفاذية التي تم الحصول عليها من هذا البئر.



شكل 1.3: موقع البئر العمودي

الإحداثيات UTM (م):

$$\begin{cases} x: & 864.769,999 \\ y: & 995.3,491,699 \end{cases}$$

لقد استخدمنا مجموعة بيانات من الجزر والتسجيلات الخاصة بالنفاذية المُحصل عليها من هذا البئر، والتي تم الحصول عليها من شركة ويذرفورد. توفر هذه البيانات النفاذية المقاسة لبناء هذا النموذج لتنبؤ النفاذية وكذلك لأغراض المقارنة في مجموع إجمالي يحتوي على 67 بيان أساسي متوفر من الجزر، تم الحصول على هذه البيانات من مقر CRD.

الجزرة 1 (3379.103401.24) أفقية
الجزرة 2 (3481.183484.05) أفقية
الجزرة 3 (3508.533513.30) أفقية

تم الحصول على هذه البيانات الأساسية من عمق 3379.10 متر إلى 3513.30 متر بسبب وجود نطاق واسع. نحن نختار بيانات غاما راي والمقاومة وطيف غاما راي والكثافة والنيوترون، ومن هذه البيانات يمكننا التنبؤ بالنفاذية.

برمجة شبكة عصبية باستخدام R

اختيار البيانات

لقد اخترنا بيانات السجلات (الدياغرافيات) التي ترتبط بالنفاذية من خلال اختيارها كمدخلات، وفي حين اخترنا بيانات النفاذية للكور كمخرجات لتقدير الانحدار الخطي للتنبؤ بالنفاذية باستخدام طريقة الشبكة العصبية.

النفاذية	
العمق (DEPT)	
أشعة غاما (GRSG)	
المقاومة (R20T)	
المقاومة (R30T)	
المقاومة (R40T)	
المقاومة (R60T)	
المقاومة (R85T)	
أشعة غاما الطيفية (GRGC)	
الكثافة (DCOR)	
النيوترون (NPRL)	

في هذا الفصل، سنقوم بتقدير الانحدار الخطي المتعدد باستخدام الشبكات العصبية للتنبؤ بالنفاذية. سنستخدم مكتبة Keras في لغة R لهذا الغرض.

تحميل المكتبات اللازمة

أول خطوة هي تحميل المكتبات اللازمة. إذا لم تكن المكتبات مثبتة، سيتم تثبيتها تلقائيًا.

```
1  # تحميل المكتبات اللازمة
2  if (!require("keras")) install.packages("keras")
3  if (!require("readxl")) install.packages("readxl")
4  if (!require("dplyr")) install.packages("dplyr")
5  if (!require("ggplot2")) install.packages("ggplot2")
6  if (!require("plotly")) install.packages("plotly")
7
8  library(keras)
9  library(readxl)
10 library(dplyr)
11 library(ggplot2)
12 library(tensorflow)
13 library(reticulate)
14 library(plotly)
15
```

شكل 2.3: تثبيت المكتبات وتحميلها

تعيين الـ Seed للحصول على نتائج متكررة

لتأكد من أن النتائج قابلة للتكرار، نقوم بتعيين الـ seed لـ R و TensorFlow و NumPy.

```
15
16  # تعيين الـ Seed لـ R, TensorFlow و NumPy للحصول على نتائج متكررة
17  set.seed(123)
18  tf$random$set_seed(123)
19  np <- import("numpy")
20  np$random$seed(as.integer(123))
21
```

شكل 3.3: الحصول على نتائج متكررة

تحميل البيانات

نقوم بتحميل البيانات من ملف Excel وتحديد الأعمدة كأعداد إذا لزم الأمر.

```
21
22  # تحميل البيانات من ملف Excel
23  file_path <- "C:/Users/SIDALI/OneDrive/Desktop/6) مجلد جديد)/putis1.xlsx"
24  data <- read_excel(file_path)
25
26  # تحويل جميع الأعمدة إلى عددية إذا كانت موجودة
27  data <- data %>%
28    mutate(across(everything(), ~as.numeric(.)))
29
```

شكل 4.3: تحميل البيانات

استيفاء القيم المفقودة

نستخدم الاستيفاء الخطي لملء القيم المفقودة في البيانات. قمنا بتعديل هذه السجلات باستخدام برنامج R.

تم ذلك من خلال الاستيفاء الخطي باستخدام القيمة الأقرب قبل وبعد. تم تسجيل قياس الانتقال في المختبر باستخدام عينة النواة مع سجل البئر الأساسي لحاسي مسعود (MD738) وسجله.

```
29
30  # استيفاء القيم المفقودة باستخدام الاستيفاء الخطي
31  interpolate_na <- function(x) {
32    if (sum(is.na(x)) > 0) {
33      x <- approx(1:length(x), x, 1:length(x), rule = 2)$y
34    }
35    return(x)
36  }
37
38  # تطبيق الاستيفاء الخطي على جميع الأعمدة
39  data[] <- lapply(data, interpolate_na)
40
```

شكل 5.3: عملية الإستيفاء الخطي

تحديد المتغيرات المستقلة والمتغير التابع

نقوم بتحديد المتغيرات المستقلة والمتغير التابع من البيانات.

```
40
41 # تحديد المتغيرات المستقلة والمتغير التابع
42 X <- data %>% select(DEPT, GR5G, R20T, R30T, R40T, R60T, R85T, GRGC, DCOR, NPRL)
43 y <- data$Permeabilité
44
```

شكل 6.3: تحديد المتغيرات المستقلة والمتغير التابع

```
44
45 # تقسيم البيانات إلى مجموعات تدريب واختبار
46 indices <- sample(1:nrow(X), size = 0.8 * nrow(X))
47 X_train <- X[indices, ]
48 y_train <- y[indices]
49 X_test <- X[-indices, ]
50 y_test <- y[-indices]
51
```

شكل 7.3: تقسيم البيانات الى بيانات تدريب وأخرى للإختبار

تطبيع البيانات

نقوم بتطبيع البيانات لضمان أن تكون موزعة بشكل مناسب للتدريب.

```
51
52 # تطبيع البيانات
53 scaler <- scale(X_train)
54 X_train_scaled <- scale(X_train)
55 X_test_scaled <- scale(X_test, center = attr(scaler, "scaled:center"), scale = attr(scaler, "scaled:scale"))
56
```

شكل 8.3: تطبيع البيانات

إنشاء نموذج الشبكة العصبية

نقوم بإنشاء نموذج شبكة عصبية تسلسلية باستخدام Keras.

```
56
57 # عدد الميزات
58 num_features <- ncol(X_train_scaled)
59
60 # إنشاء نموذج شبكة عصبية
61 model <- keras_model_sequential() %>%
62   layer_dense(units = 64, activation = 'relu', input_shape = c(num_features)) %>%
63   layer_dense(units = 128, activation = 'relu', kernel_initializer = 'he_normal') %>%
64   layer_dense(units = 1, activation = 'linear')
65
```

شكل 9.3: إنشاء نموذج شبكة عصبية

ترجمة النموذج

نقوم بترجمة النموذج باستخدام دالة الخسارة والمُحسِّن المناسبين.

تدريب النموذج

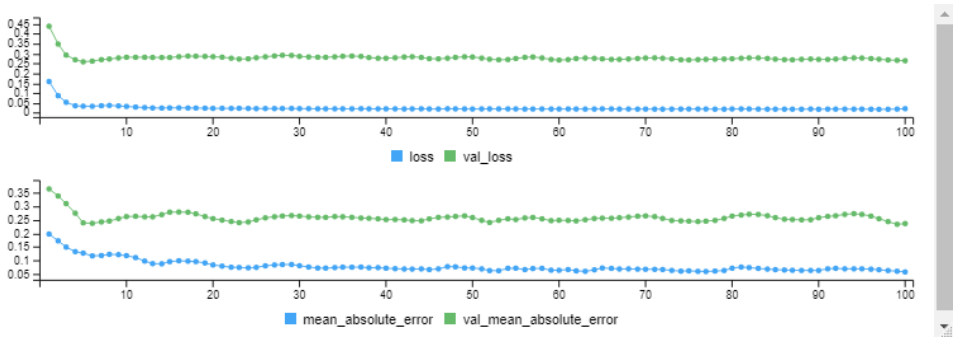
نقوم بتدريب النموذج على مجموعة البيانات التدريبية.

```
65
66  # ترجمة النموذج
67  model %>% compile(
68    loss = 'mean_squared_error',
69    optimizer = optimizer_adam(),
70    metrics = c('mean_absolute_error')
71  )
72
```

شكل 10.3: ترجمة النموذج

```
72
73  # تدريب النموذج
74  history <- model %>% fit(
75    X_train_scaled, y_train,
76    epochs = 100,
77    batch_size = 32,
78    validation_split = 0.2
79  )
80
```

شكل 11.3: تدريب النموذج



شكل 12.3: مسار التطور أثناء التدريب

تقييم النموذج

نقوم بتقييم النموذج على مجموعة البيانات الاختبارية.

```
80
81  # تقييم النموذج
82  loss <- model %>% evaluate(X_test_scaled, y_test)
83  cat('Test Loss:', loss, '\n')
84
```

شكل 13.3: تقييم النموذج

```
> # تقييم النموذج
> loss <- model %>% evaluate(X_test_scaled, y_test)

1/1 [==============================] - ETA: 0s - loss: 0.0173 - mean_absolute_error: 0.0820
1/1 [==============================] - 2s 2s/step - loss: 0.0173 - mean_absolute_error: 0.0820
> cat('Test Loss:', loss, '\n')
Test Loss: 0.01725167 0.0820327
```

شكل 14.3: نتائج التقييم

إجراء التنبؤات

نقوم بإجراء التنبؤات باستخدام النموذج المدرب.

```
84
85  # إجراء التنبؤات
86  predictions <- model %>% predict(X_test_scaled)
87
```

شكل 15.3: إجراء تنبؤات

حساب مقاييس الأداء

نقوم بحساب معامل التحديد R^2، والخطأ المطلق الوسطي MAE، والجذر التربيعي لمتوسط مربع الخطأ $RMSE$، ومتوسط مربع الخطأ MSE.

```
87
88  # حساب معامل التحديد R-squared والخطأ المطلق الوسطي MAE
89  r_squared <- cor(y_test, predictions)^2
90  mae <- mean(abs(y_test - predictions))
91
92  # حساب MSE و RMSE
93  mse <- mean((y_test - predictions)^2)
94  rmse <- sqrt(mse)
95
96  cat('R-squared:', r_squared, '\n')
97  cat('Mean Absolute Error:', mae, '\n')
98  cat('Mean Squared Error:', mse, '\n')
99  cat('Root Mean Squared Error:', rmse, '\n')
100
```

شكل 16.3: حساب مقاييس الأداء

نحصل على هذه النتائج :

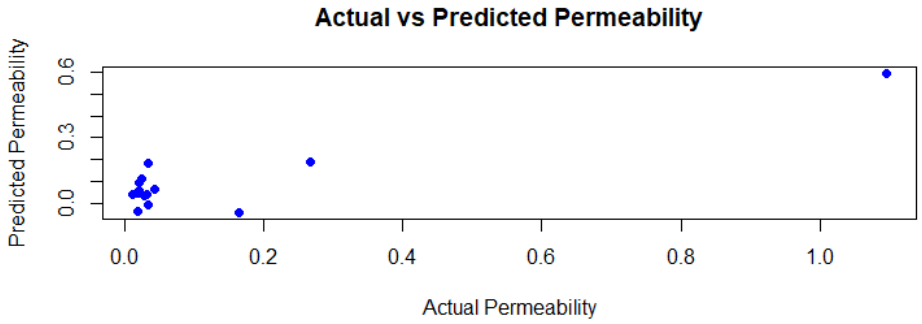
$$\begin{cases} R-squared = 0.8872093 \\ MSE = 0.01725167 \\ MAE = 0.0820327 \\ RMSE = 0.1313456 \end{cases}$$

رسم النتائج

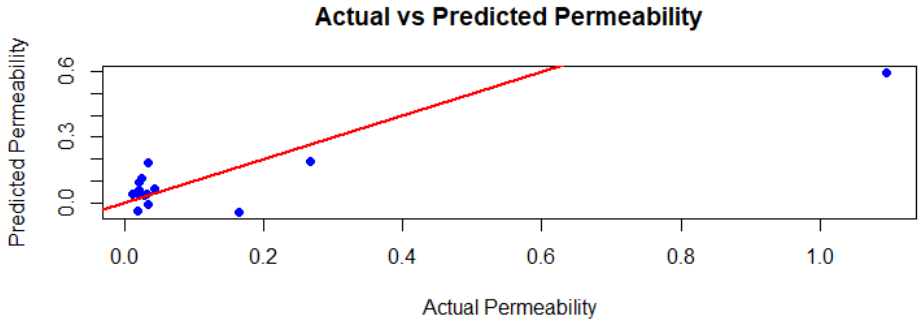
نقوم برسم القيم الفعلية مقابل القيم المتنبأة ومسار التطور خلال التدريب.

```
100
101 # رسم القيم الفعلية مقابل القيم المتنبأة
102 plot(y_test, predictions, main = "Actual vs Predicted Permeability",
103      xlab = "Actual Permeability", ylab = "Predicted Permeability", pch = 19, col = "blue")
104 abline(0, 1, col = "red", lwd = 2)
105
106 # رسم مسار التطور خلال التدريب
107 plot(history$metrics$val_loss, type = 'l', col = 'red', ylim = range(c(history$metrics$loss,
108                                                                        history$metrics$val_loss)),
109      ylab = 'Loss', xlab = 'Epoch', main = "Training and Validation Loss")
110 lines(history$metrics$loss, col = 'blue')
111 legend("topright", legend = c("Training Loss", "Validation Loss"), col = c("blue", "red"), lty = 1)
112
```

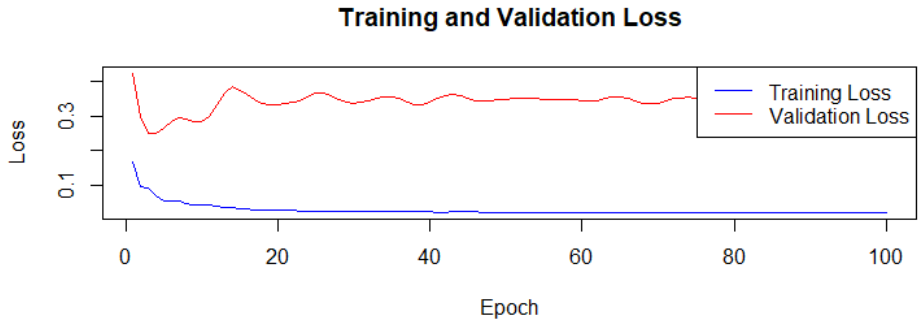
شكل 17.3: رسم القيم الفعلية مقابل القيم المتنبأة ومسار التطور خلال التدريب.



شكل 18.3: رسم القيم الفعلية مقابل القيم المتنبأ بها



شكل 19.3: رسم خط الإنحدار

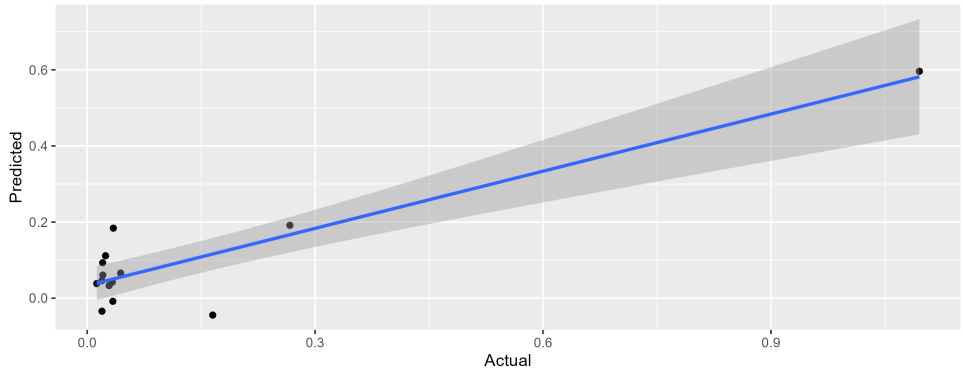


شكل 20.3: رسم مسار التطور أثناء التدريب

```
g <- ggplot(data.frame(Actual = y_test, Predicted = as.vector(predictions)),
            aes(x = Actual, y = Predicted)) +
  geom_point() +
  geom_smooth(method = 'lm')
```

شكل 21.3: رسم القيم الفعلية مقابل القيم المتنبأ بها بصيغة أفضل

تحليل وتفسير النتائج



شكل 22.3: رسم القيم الفعلية مقابل القيم المتنبأ بها بصيغة أفضل

الصورة المدرجة هي لرسم يوضح العلاقة بين القيم الفعلية والقيم المتنبأ بها للنموذج. النقاط السوداء تمثل القيم الفعلية مقابل القيم المتنبأ بها، بينما الخط الأزرق يمثل خط الانحدار الذي يحاول التقريب بين القيم الفعلية والمتنبأ بها.

تحليل الرسم

- **توزيع النقاط:** معظم النقاط متركزة بالقرب من الجزء السفلي الأيسر من الرسم، مما يشير إلى أن معظم القيم الفعلية والمتنبأ بها صغيرة. هناك تشتت واضح في النقاط كلما ابتعدنا عن الصفر، مما يشير إلى وجود بعض الأخطاء في التنبؤ بالقيم العالية للنفاذية.

- **خط الانحدار:** الخط الأزرق يمثل خط الانحدار الذي يحاول تقريب العلاقة بين القيم الفعلية والمتنبأ بها. إذا كان النموذج دقيقًا تمامًا، فإن جميع النقاط السوداء ستقع على هذا الخط. الانحرافات عن الخط الأزرق تشير إلى الأخطاء في التنبؤ، وكلما زادت هذه الانحرافات، زادت الأخطاء.

- **منطقة الثقة:** المنطقة الرمادية حول الخط الأزرق تمثل منطقة الثقة، وهي تعكس درجة الثقة في تقديرات النموذج. كلما كانت المنطقة أضيق، كانت الثقة أعلى في التنبؤات.

التفسير

هذه النتائج تشير إلى أن النموذج لديه قدرة جيدة على التنبؤ بالقيم الفعلية للنفاذية، خاصة للقيم الصغيرة. ومع ذلك، هناك بعض التشتت والأخطاء في التنبؤ بالقيم العالية. يمكن تحسين النموذج من خلال تحسين المعايير أو استخدام نماذج أكثر تعقيدًا لتقليل الأخطاء وزيادة دقة التنبؤات.

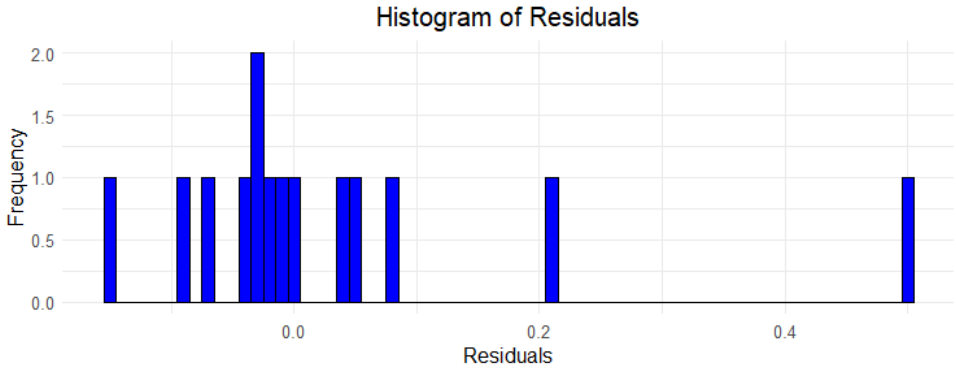
رسم البواقي

نقوم بحساب ورسم البواقي.

```
112
113 # حساب البواقي
114 residuals <- y_test - predictions
115
116 # رسم المدرج التكراري للبواقي بدقة أعلى
117 ggplot(data.frame(Residuals = residuals), aes(x = Residuals)) +
118   geom_histogram(binwidth = 0.01, fill = 'blue', color = 'black') +
119   ggtitle("Histogram of Residuals") +
120   xlab("Residuals") +
121   ylab("Frequency") +
122   theme_minimal() +
123   theme(
124     plot.title = element_text(hjust = 0.5, size = 15),
125     axis.title = element_text(size = 12),
126     axis.text = element_text(size = 10)
127   )
128
129 # رسم البواقي مقابل القيم المتوقعة بتفاصيل إضافية
130 ggplot(data.frame(Predicted = as.vector(predictions), Residuals = residuals),
131        aes(x = Predicted, y = Residuals)) +
132   geom_point(color = 'blue', alpha = 0.5) +
133   geom_smooth(method = 'loess', color = 'red', linetype = 'dashed') +
134   geom_hline(yintercept = 0, color = 'red', linetype = 'dashed') +
135   ggtitle("Residuals vs Predicted") +
136   xlab("Predicted Permeability") +
137   ylab("Residuals") +
138   theme_minimal() +
139   theme(
140     plot.title = element_text(hjust = 0.5, size = 15),
141     axis.title = element_text(size = 12),
142     axis.text = element_text(size = 10)
143   )
144
```

شكل 23.3: حساب ورسم البواقي

تحليل وتفسير النتائج



شكل 24.3: رسم المدرج التكراري للبواقي

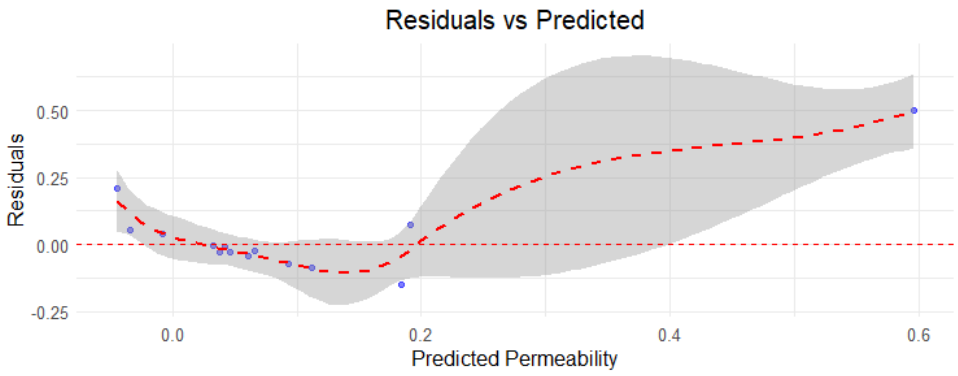
في الرسم البياني المدرج التكراري للبواقي، يمكننا تحليل النتائج كالتالي:

1. **التوزيع**:
   - يبدو أن معظم البواقي تتركز حول الصفر، مما يشير إلى أن التنبؤات عادة ما تكون قريبة من القيم الفعلية.
2. **التماثل**:
   - المدرج التكراري يظهر بعض التماثل حول الصفر، مما يدل على أن الأخطاء في التنبؤ تتوزع بشكل متساوٍ إلى حد ما بين القيم السالبة والموجبة.
   - البواقي الإيجابية والسلبية تشير إلى وجود حالات تنبأ فيها النموذج بقيم أعلى أو أقل من القيم الفعلية، على التوالي.
3. **التكرار**:
   - نلاحظ أن هناك بواقي متكررة عند القيمة الصفر، مما يشير إلى أن هناك عدة تنبؤات كانت مطابقة تقريباً للقيم الفعلية.
   - أيضاً، هناك بعض القيم المتطرفة (outliers) التي تظهر عند الأطراف، والتي قد تحتاج إلى مزيد من التحقيق.
4. **الانحراف**:
   - تركز البواقي حول الصفر يعني أن النموذج لا يعاني من انحياز كبير، حيث أن الأخطاء موزعة بشكل متساوٍ تقريباً حول الصفر.

خلاصة

الرسم البياني يظهر أن النموذج أداءه جيد بشكل عام مع بعض الأخطاء الكبيرة. التركيز على تحسين أداء النموذج على هذه النقاط المتطرفة قد يساعد في تحسين الدقة العامة للنموذج. القيم الإحصائية الأخرى مثل R^2 و MAE و MSE و RMSE تدعم هذا الاستنتاج، حيث أن R^2 يظهر دقة عالية للنموذج، ولكن قيم الخطأ تشير إلى وجود بعض التباين في الدقة.

تحليل وتفسير النتائج



شكل 25.3: رسم البواقي مقابل القيم المتوقعة

الصورة المرفقة هي لرسم يوضح البواقي مقابل القيم المتوقعة مع إضافة تفاصيل توضيحية. الرسم يحتوي على النقاط الزرقاء التي تمثل الفرق بين القيم الفعلية والمتنبأ بها (البواقي)، والخط الأحمر المنقط يمثل اتجاه العلاقة بين البواقي والقيم المتنبأ بها. يتم استخدام الخط الأفقي الأحمر المنقط عند الصفر كمرجع.

تحليل الرسم

- **التوزيع العام للبواقي:** يتمحور معظم البواقي بالقرب من الصفر، مما يشير إلى أن الأخطاء في التنبؤ ليست كبيرة في معظم الحالات. هناك تشتت واضح للبواقي عندما تكون القيم المتنبأ بها أكبر، مما يشير إلى أن النموذج قد يواجه صعوبة في التنبؤ بدقة للقيم العالية للنفاذية.
- **اتجاه العلاقة:** الخط الأحمر المنقط المائل يشير إلى عدم وجود نمط في الأخطاء. في الحالات المثالية، يجب أن تكون البواقي موزعة بشكل عشوائي حول الصفر دون أي نمط واضح. الانحناء في الخط الأحمر يشير إلى أن هناك علاقة غير خطية بين القيم المتنبأ بها والبواقي.
- **النقاط البعيدة:** يمكن ملاحظة وجود بعض النقاط البعيدة عن الخط الأفقي الأحمر، مما يشير إلى وجود بعض القيم التي لم يتم التنبؤ بها بدقة كبيرة. هذه النقاط يمكن أن تكون ناتجة عن بيانات شاذة أو متغيرات مؤثرة لم يتم تضمينها في النموذج.

التفسير

هذه النتائج تشير إلى أن النموذج لديه قدرة جيدة على التنبؤ بالنفاذية، ولكن هناك بعض القيم التي لم يتم التنبؤ بها بدقة. يمكن تحسين النموذج من خلال النظر في متغيرات إضافية أو استخدام نماذج أكثر تعقيدًا لالتقاط العلاقات غير الخطية بين المتغيرات.

استنتاج

من خلال هذا البحث، استطعنا أن نثبت بأن الانحدار الخطي المتعدد هو أداة قوية للتنبؤ والتحليل في العديد من المجالات. النتائج التي حصلنا عليها تدعم بقوة الفرضية القائلة بأن متغيرات مثل العمق وأشعة غاما والمقاومات وأشعة غاما الطيفية والكثافة والنيوترون لها تأثير مباشر ومعنوي على النفاذية. هذا يظهر ليس فقط القدرة التنبؤية للنموذج، ولكن أيضاً أهميته في تطبيقات عملية مثل التخطيط لعمليات الحفر وتقييم المخاطر الجيولوجية.

كما أن الاعتماد على طريقة الشبكات العصبية لتقدير معالم الانحدار قد فتح الباب أمام استكشاف أساليب جديدة في البحث العلمي القائم على البيانات، وهو ما يمكن أن يؤدي إلى تحسينات ملحوظة في الدقة والكفاءة في الدراسات المستقبلية.

وفي الختام، يمكن القول إن هذه الدراسة لم تسهم فقط في تعزيز فهمنا للعلاقات بين المتغيرات في نماذج الانحدار المتعدد، بل قدمت أيضاً مسارات للتطوير المستقبلي في البحوث ذات الصلة، مما يعكس الإمكانات الواسعة لاستخدام تقنيات التحليل الكمي في تعميق معارفنا وتحسين تقنياتنا.

خاتمة

ختاماً، تجدر الإشارة إلى أننا عبر هذا البحث قد قمنا بجمع وتفسير وتحليل أهم المعلومات الخاصة بموضوع الانحدار الخطي المتعدد، وتقدير معالمه باستخدام طريقة الشبكات العصبية. وفي الأخير، قمنا بإعطاء مثال وتحليله باستخدام لغة البرمجة الإحصائية R من أجل التنبؤ بالنفاذية. تم استخدام نموذج الانحدار الخطي المتعدد، والذي اعتبرت فيه متغيرات العمق وأشعة غاما والمقاومات وأشعة غاما الطيفية والكثافة والنيوترون كمتغيرات تفسيرية ومتغير النفاذية كمتغير تابع. أظهرت نتائج نموذج الانحدار أن هناك علاقة معنوية بين المتغيرات المستقلة والمتغير التابع.

من خلال التحليل الذي قمنا به باستخدام الشبكات العصبية، حصلنا على النتائج التالية:

$$\begin{cases} R-squared = 0.8872093 \\ MSE = 0.01725167 \\ MAE = 0.0820327 \\ RMSE = 0.1313456 \end{cases}$$

- **R-squared**: تشير إلى أن النموذج يفسر نسبة كبيرة من التباين في البيانات، مما يدل على قوة التنبؤ للنموذج.
- **(MAE)**: يشير إلى أن متوسط الخطأ المطلق للتنبؤات، مما يدل على دقة عالية.
- **(MSE)**: يشير إلى أن متوسط مربع الخطأ، وهو منخفض ويعكس دقة النموذج.
- **(RMSE)**: يشير إلى أن جذر متوسط مربع الخطأ، مما يدل على أن الأخطاء في التنبؤات صغيرة نسبيًا.

تفتح هذه النتائج آفاقًا جديدة لاستخدام الشبكات العصبية في تقدير النماذج التنبؤية في مجالات متعددة بما في ذلك علوم الأرض والبيئة. يوصى بإجراء المزيد من الدراسات لفحص تأثير متغيرات أخرى قد تؤثر في النفاذية مثل الضغط ودرجة الحرارة، للتوصل إلى نماذج أكثر دقة وشمولية.

أتمنى من الله أن تكون كلماتي نالت إعجابكم، بتوفيق منه تعالى تمكنت من اجتياز جملة من العراقيل والصعوبات لإتمام هذا العمل والذي لا يخلو من النقائص فلله الكمال والتمام.

المصادر

المصادر العربية

[1] رشا طلال عطاء، آلاء شمس الله نور الله الخزعل، علي ياسين غني، استعمال الشبكات العصبية الاصطناعية في تحسين فاعلية القوائم المالية في مصرف الشرق الأوسط، مجلة المستنصرية للدراسات العربية والدولية، العراق، المجلد 6، العدد 3، ص 45-67، 2020.

[2] بن نور فريد، نايت مرزوق محمد العربي، استخدام نماذج الذكاء الصناعي للتنبؤ باحتياطات الصرف الأجنبي في الجزائر، مجلة اقتصاديات شمال افريقيا، المجلد 15، العدد 2، 2019.

[3] فاطمة بوادو، التنبؤ بمبيعات المؤسسات الجزائرية باستخدام نماذج السلاسل الزمنية وتقنية الشبكات العصبية الاصطناعية-دراسة حالة مؤسسة سونلغاز بالشلف، أطروحة دكتوراه، جامعة ابن خلدون تيارت، 2016.

[4] هدى خليل إبراهيم الحسيني، علي محمد ثجيل المعموري، استخدام الشبكات العصبية الاصطناعية في تطوير دور مراقب الحسابات في اكتشاف الأخطاء الجوهرية: بحث تطبيقي في الشركة العامة للصناعات الكهربائية وشركة نصر العامة للصناعات الميكانيكية، مجلة دراسات محاسبية ومالية، جامعة بغداد، العراق، المجلد 10، العدد 86، ص 45-68، 2016.

[5] عبد الفتاح مصطفى محمد، الانحدار المتعدد، جامعة مصر، كلية العلوم، 2020.

[6] وحيد محمود رمو، التنقيب المحاسبي عن البيانات باستخدام الشبكات العصبية: دراسة حالة، مجلة الاقتصاد وعلوم الإدارة، العراق، العدد 666، المجلد 9، ص 12-34، 2016.

[7] تومي صالح، مدخل لنظرية القياس الاقتصادي، ديوان المطبوعات الجامعية، الجزائر، 1999.

المصادر الأجنبية

[8] Christopher M. Bishop, *Neural Networks for Pattern Recognition*, Oxford University Press, 1995.

[9] François Chollet, *Deep Learning with Python*, Manning Publications, 2017.

[10] Aurélien Géron, *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow*, 2nd Edition, O'Reilly Media, 2019.

[11] Xavier Glorot and Yoshua Bengio, "Understanding the difficulty of training deep feedforward neural networks." In *Proceedings of the thirteenth international conference on artificial intelligence and statistics*, pages 249-256, 2010.

[12] Ian Goodfellow, Yoshua Bengio, and Aaron Courville, *Deep Learning*, MIT Press, 2016.

[13] Trevor Hastie, Robert Tibshirani, and Jerome Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd Edition, Springer Series in Statistics, Springer, 2009.

[14] Simon Haykin, *Neural Networks and Learning Machines*, 3rd Edition, Prentice Hall, 2009.

[15] Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun, "Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification," *IEEE International Conference on Computer Vision (ICCV)*, 2015, pp. 1026-1034.

[16] Geoffrey Hinton, Nitish Srivastava, and Kevin Swersky, "Lecture 6a Overview of mini-batch gradient descent," *Coursera, Neural Networks for Machine Learning*, 2012.

[17] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton, "Deep learning." *Nature*, 521(7553):436-444, 2015.

[18] S. Marsland, *Machine Learning: An Algorithmic Perspective, Second Edition*, Chapman and Hall/CRC, 2014.

[19] Michael Nielsen, *Neural Networks and Deep Learning*, Determination Press, 2015.

[20] Jorge Nocedal, Stephen J. Wright, *Numerical Optimization*, Springer, Second Edition, 2006.

[21] S. Raschka & V. Mirjalili, *Python Machine Learning*, Packt Publishing, 2017.

[22] Matthew D. Zeiler and Rob Fergus, "Visualizing and understanding convolutional networks." In *European conference on computer vision*, pages 818-833, 2014.

[23] Chiyuan Zhang, Samy Bengio, Moritz Hardt, Benjamin Recht, and Oriol Vinyals, "Understanding deep learning requires rethinking generalization." *arXiv preprint arXiv:1611.03530*, 2017.

ملخص

هذه المذكرة تستعرض كيفية تحسين تقديرات الانحدار الخطي المتعدد باستخدام الشبكات العصبية من خلال برمجة لغة R. الانحدار الخطي، الذي يُعد إجراء إحصائياً معتمداً على نطاق واسع للتنبؤ العلمي والدقيق، تم تحليله هنا بطريقة مبتكرة من خلال تطبيق الشبكات العصبية لتقدير العلاقات بين المتغيرات المستقلة والمتغير التابع.

تركز المذكرة على استخدام الشبكات العصبية كطريقة وحيدة لتحليل البيانات، حيث تم استبدال الطرق التقليدية مثل المربعات الصغرى بتقنيات تعلم الآلة التي توفر مرونة أكبر ودقة محسنة في التقدير. من خلال برمجة لغة R، تم تطوير نماذج الشبكات العصبية لتفسير البيانات المعقدة وتحديد العوامل الأساسية التي تؤثر على المتغير التابع.

تُظهر نتائج هذه المذكرة كيف أن استخدام الشبكات العصبية يمكن أن يعزز بشكل كبير من دقة التنبؤات في نماذج الانحدار الخطي المتعدد، ويفتح آفاقًا جديدة لاستكشاف العلاقات المعقدة في البيانات. من خلال هذه الدراسة، نقدم مسارًا للتطوير المستقبلي في الأبحاث، مع التأكيد على إمكانيات استخدام تقنيات التحليل الكمي المتقدمة في تحسين فهمنا للظواهر المختلفة.

الكلمات المفتاحية: الانحدار الخطي المتعدد، الشبكات العصبية، تعلم الآلة، برمجة R، التنبؤ الإحصائي

Summary

This thesis explores how to improve multiple linear regression estimates using neural networks through the R programming language. Linear regression, a widely recognized statistical method for scientific and accurate forecasting, has been innovatively analyzed here by applying neural networks to estimate relationships between independent variables and the dependent variable.

The thesis focuses on using neural networks as the sole method for data analysis, replacing traditional methods such as least squares with machine learning techniques that offer greater flexibility and improved accuracy in estimation. Through R programming, neural network models have been developed to interpret complex data and identify key factors influencing the dependent variable.

The results of this thesis demonstrate how the use of neural networks can significantly enhance prediction accuracy in multiple linear regression models, opening new avenues for exploring complex relationships in data. Through this study, we present a pathway for future research development, emphasizing the potential of using advanced quantitative analysis techniques to enhance our understanding of various phenomena.

Keywords: Multiple Linear Regression, Neural Networks, Machine Learning, R Programming, Statistical Prediction