الكلية التطبيقية

الجامعة السعودية الإلكترونية

SAUDI ELECTRONIC UNIVERSITY

2011-1432

دبلوم الذكاء
الاصطناعي التوليدي

اسم المقرر: الإحصاء و تحليل البيانات لعلوم الآلة
رمز المقرر:STA104
الموضوع: تقييم النموذج R^2، MAE، MSE

الموضوع الثاني عشر: تقييم النموذج R^2، MAE، MSE

Restricted - مقيد

الأهداف

في نهاية هذا الدرس ستكون قادرًا على أن:

- ✓ إدراك الفرق بين الخطأ Error والبواقي Residualsفي سياق نماذج الانحدار
- ✓ فهم معيار متوسط الخطأ المطلق MAE وكيفية حسابه وتفسيره
- ✓ فهم معيار متوسط مربع الخطأ MSE وخصائصه الرياضية
- ✓ فهم معيار جذر متوسط مربع الخطأ RMSE ومتى يُستخدم
- ✓ فهم معامل التحديد R^2 وتفسيره الإحصائي العميق
- ✓ المقارنة بين المعايير المختلفة ومعرفة متى يُستخدم كل منها
- ✓ فهم مفهوم مجموع المربعات Sum of Squares بأنواعه المختلفة
- ✓ القدرة على تفسير قيم هذه المعايير في السياق العملي
- ✓ إدراك حدود ومحاذير استخدام كل معيار
- ✓ فهم العلاقة بين هذه المعايير والافتراضات الإحصائية للانحدار
- ✓ التمييز بين التقييم على البيانات التدريبية والاختبارية

وصف الدرس

يُعد تقييم نماذج الانحدار خطوة حاسمة في عملية بناء نماذج تعلم الآلة، حيث لا يكفي بناء نموذج دون معرفة مدى جودة أدائه وقدرته على التنبؤ بدقة. في هذا الدرس، سنتعمق في فهم أهم المعايير الإحصائية المستخدمة لتقييم نماذج الانحدار، وهي متوسط الخطأ المطلق Mean Absolute Error - MAE، متوسط مربع الخطأ Mean Squared Error - MSE، جذر متوسط مربع الخطأ Root Mean Squared Error - RMSE، ومعامل التحديد Coefficient of Determination - R^2. سنبدأ بفهم المنطق الرياضي والإحصائي وراء كل معيار، ولماذا صُمم بهذه الطريقة المحددة، وما الذي يقيسه بالضبط. سنستكشف كيف تتعامل كل من هذه المعايير مع الأخطاء الكبيرة مقابل الصغيرة، ولماذا يعطي بعضها وزناً أكبر للقيم الشاذة. سنتناول أيضاً الأساس النظري لمعامل R^2 وعلاقته بمفهوم تباين البيانات، وكيف يمكن تفسيره كنسبة التباين المفسر من قبل النموذج. سنناقش متى يكون كل معيار أكثر ملاءمة من الآخر، ونستعرض أمثلة توضيحية تبين كيف يمكن لنماذج مختلفة أن تحصل على قيم مختلفة لهذه المعايير. كما سنتطرق إلى القيود والمحاذير المهمة لكل معيار، مثل حساسية MSE للقيم الشاذة، وإمكانية حصول R^2 على قيم سالبة في حالات معينة. الهدف النهائي هو تزويد الطالب بفهم عميق وشامل يمكنه من اختيار المعيار المناسب لتقييم نماذجه وتفسير النتائج بشكل صحيح ودقيق.

عنوان | لماذا نحتاج لتقييم النماذج؟

الحاجة الأساسية:

عندما نبني نموذج انحدار للتنبؤ بقيم مستمرة، لا يكفي أن نقول "النموذج جاهز". نحتاج لطريقة موضوعية وكمية لقياس مدى جودة هذا النموذج ومدى دقة تنبؤاته. بدون معايير تقييم واضحة، لا نستطيع:

1. مقارنة النماذج المختلفة:

لو أنشأنا عدة نماذج بخوارزميات مختلفة أو بمعاملات مختلفة، كيف نختار الأفضل؟ نحتاج لمعيار موحد للمقارنة.

2. معرفة جودة النموذج المطلقة:

حتى لو كان لدينا نموذج واحد، نحتاج لمعرفة: هل هو جيد بما يكفي لاستخدامه في الواقع؟ أم أن دقته متدنية ولا يمكن الاعتماد عليه؟

3. تحديد مجالات التحسين:

معايير التقييم تساعدنا في فهم أين يخطئ النموذج وبأي مقدار، مما يوجهنا نحو طرق تحسينه.

4. اتخاذ قرارات مبنية على البيانات:

في البيئات العملية، القرارات الاستثمارية أو التشغيلية تُبنى على دقة النموذج. معايير التقييم تعطينا ثقة (أو تحذير) حول الاعتماد على النموذج.

5. التواصل مع الآخرين:

المعايير الكمية لغة مشتركة نستخدمها لشرح جودة النموذج للمدراء، العملاء، أو الزملاء.

Restricted - مقيد

عنوان | الخطأ والبواقي - المفاهيم الأساسية

قبل البدء في المعايير، نحتاج لفهم:

القيمة الفعلية: y_i - (Actual Value)

القيمة الحقيقية للمتغير التابع في البيانات. هذا ما نعرفه من الواقع أو القياسات الفعلية.

القيمة المتوقعة: $\hat{y}_i$ - (Predicted Value)

القيمة التي يتنبأ بها النموذج للمتغير التابع. هذا ما يخرجه نموذجنا بناءً على المتغيرات المستقلة.

الخطأ أو البواقي: e_i - (Error/Residual)

الفرق بين القيمة الفعلية والمتوقعة:

$$e_i = y_i - \hat{y}_i$$

أنواع الأخطاء:

- **خطأ موجب: ($e_i > 0$)** النموذج توقع قيمة أقل من الواقع(Underestimation)
- **خطأ سالب: ($e_i < 0$)** النموذج توقع قيمة أكبر من الواقع(Overestimation)
- **خطأ صفري: ($e_i = 0$)** النموذج توقع بدقة تامة (نادر جداً)

مثال توضيحي:

- السعر الفعلي لمنزل: 500,000 ريال(y_i)
- السعر المتوقع من النموذج: 480,000 ريال($\hat{y}_i$)
- الخطأ: 500,000 - 480,000 = 20,000 ريال(e_i)
- التفسير: النموذج قلل من تقدير السعر بمقدار 20,000 ريال

ملاحظة أساسية:

كل معايير التقييم التي سندرسها تعتمد على هذه الأخطاء، لكن كل منها يتعامل معها بطريقة مختلفة.

Restricted - مقيد

عنوان | لماذا لا نستخدم متوسط الخطأ فقط؟

السؤال الطبيعي:
لماذا لا نحسب ببساطة متوسط الأخطاء ونستخدمه كمعيار؟
متوسط الخطأ :Mean Error
$ME = (1/n)\ \Sigma(y_i - \hat{y}_i)$
المشكلة الكبرى - التعويض والإلغاء:
الأخطاء الموجبة والسالبة تلغي بعضها البعض! النموذج قد يكون سيئاً جداً لكن يبدو جيداً بهذا المقياس.
مثال توضيحي:
لنفترض نموذجاً للتنبؤ بدرجات الطلاب في امتحان (من 100):
- الطالب 1: الفعلي 90، المتوقع 70، الخطأ = +20
- الطالب 2: الفعلي 60، المتوقع 80، الخطأ = -20
- الطالب 3: الفعلي 75، المتوقع 65، الخطأ = +10
- الطالب 4: الفعلي 55، المتوقع 65، الخطأ = -10

حساب متوسط الخطأ:
$ME = (20 + (-20) + 10 + (-10)) / 4 = 0 / 4 = 0$
النتيجة المضللة:
متوسط الخطأ = صفر! يبدو وكأن النموذج مثالي، لكن في الواقع أخطأ في كل تنبؤ بمقادير كبيرة (10-20 درجة)!
الدرس المستفاد:
نحتاج لمعايير تأخذ في الاعتبار حجم الخطأ بغض النظر عن اتجاهه (موجب أو سالب). هذا ما تفعله المعايير MAEو MSEو RMSE.

Restricted - مقيد

عنوان | متوسط الخطأ المطلق - MAE

Mean Absolute Error (MAE):

التعريف الرياضي:

$MAE = (1/n) \Sigma|y_i - \hat{y}_i|$

حيث:

- n = عدد الملاحظات (العينات)
- y_i = القيمة الفعلية للملاحظة i
- $\hat{y}_i$ = القيمة المتوقعة للملاحظة i
- $|...|$ = القيمة المطلقة (تحول كل قيمة إلى موجبة)

الفكرة الأساسية:

MAE هو متوسط القيم المطلقة للأخطاء. بأخذ القيمة المطلقة، نتجنب مشكلة الإلغاء ونركز على "حجم" الخطأ فقط بغض النظر عن اتجاهه.

خطوات الحساب:

1. احسب الخطأ لكل ملاحظة: $e_i = y_i - \hat{y}_i$
2. خذ القيمة المطلقة لكل خطأ: $|e_i|$
3. اجمع كل القيم المطلقة
4. اقسم على عدد الملاحظات

خصائص MAE:

- دائماً قيمة موجبة أو صفر
- الوحدة نفس وحدة المتغير التابع (مثلاً إذا كنا نتنبأ بالأسعار بالريال، MAE بالريال)
- سهل الفهم والتفسير
- معيار قوي - أقل حساسية للقيم الشاذة

التفسير:

إذا كان MAE = 15,000 ريال في نموذج التنبؤ بأسعار المنازل، فهذا يعني أن النموذج في المتوسط يخطئ بمقدار 15,000 ريال (أعلى أو أقل من السعر الفعلي).

Restricted - مقيد

عنوان | MAE - مثال عملي تفصيلي

مثال: التنبؤ بدرجات الحرارة اليومية (بالدرجة المئوية)

البيانات: اليوم	الحرارة الفعلية (y_i)	الحرارة المتوقعة($\hat{y}_i$)	الخطأ (e_i)	القيمة المطلقة $\|e_i\|$
1	25	23	2	2
2	28	30	-2	2
3	22	24	-2	2
4	30	28	2	2
5	26	26	0	0

الحساب خطوة بخطوة:

الخطوة 1: حساب الأخطاء (موجودة في الجدول أعلاه)

الخطوة 2: حساب القيم المطلقة (موجودة في الجدول)

الخطوة 3: جمع القيم المطلقة:

$\Sigma|e_i| = 2 + 2 + 2 + 2 + 0 = 8$

الخطوة 4: القسمة على عدد الملاحظات:

MAE = 8 / 5 = 1.6 درجة مئوية

التفسير:

النموذج في المتوسط يخطئ بمقدار 1.6 درجة مئوية في التنبؤ بدرجة الحرارة. هذا يعني أن توقعاتنا عادة ما تكون ضمن نطاق ±1.6 درجة من القيمة الفعلية.

ملاحظة:

لو حسبنا متوسط الخطأ العادي (بدون القيمة المطلقة):

ME = (2 + (-2) + (-2) + 2 + 0) / 5 = 0

سيكون صفراً ويوحي خطأً بأن النموذج مثالي!

Restricted - مقيد

عنوان | MAE - المزايا والعيوب

مزايا MAE:

1. سهولة التفسير:

MAE له تفسير بديهي ومباشر: "متوسط الخطأ المطلق في وحدات المتغير الأصلي". أي شخص يمكنه فهمه حتى بدون خلفية إحصائية عميقة.

2. نفس وحدة المتغير:

إذا كنا نتنبأ بالأسعار بالريال، MAE بالريال. إذا كنا نتنبأ بالمسافة بالكيلومتر، MAE بالكيلومتر. هذا يجعل التفسير طبيعياً.

3. قوي ضد القيم الشاذة (Robust):

لأن MAE يستخدم القيمة المطلقة (وليس التربيع)، فإن القيم الشاذة الكبيرة لا تؤثر عليه بشكل كبير جداً. كل خطأ يساهم بشكل خطي (مباشر) في النتيجة النهائية.

4. معاملة متساوية للأخطاء:

كل خطأ يُعامل بالتساوي بغض النظر عن حجمه. خطأ بمقدار 10 يساهم بـ 10، وخطأ بمقدار 100 يساهم بـ 100 (خطياً).

عيوب MAE:

1. عدم التمييز بين الأخطاء الكبيرة والصغيرة:

في بعض التطبيقات، الأخطاء الكبيرة أخطر بكثير من الصغيرة، لكن MAE يعاملها بالتساوي. مثلاً، في التشخيص الطبي، خطأ كبير قد يكون كارثياً.

2. غير قابل للاشتقاق عند الصفر:

من الناحية الرياضية، دالة القيمة المطلقة غير قابلة للاشتقاق عند نقطة الصفر، مما قد يسبب صعوبات في بعض خوارزميات التحسين (Optimization).

3. لا يعطي وزناً إضافياً للقيم الشاذة:

في حين أن هذه ميزة في بعض الحالات، قد تكون عيباً في حالات أخرى حيث نريد معاقبة الأخطاء الكبيرة بشدة أكبر.

Restricted - مقيد

عنوان | متوسط مربع الخطأ MSE

Mean Squared Error (MSE):

التعريف الرياضي:

$$MSE = (1/n)\ \Sigma(y_i - \hat{y}_i)^2$$

حيث:

- n = عدد الملاحظات
- y_i = القيمة الفعلية للملاحظة i
- $\hat{y}_i$ = القيمة المتوقعة للملاحظة i
- $(\ldots)^2$ = التربيع

الفكرة الأساسية:

MSE هو متوسط مربعات الأخطاء. بدلاً من أخذ القيمة المطلقة (كما في MAE)، نربّع كل خطأ. التربيع يحقق هدفين:

1. يحول كل القيم إلى موجبة (فيتجنب مشكلة الإلغاء)
2. يعطي وزناً أكبر للأخطاء الكبيرة (عقوبة تزداد بشكل تربيعي)

خطوات الحساب:

1. احسب الخطأ لكل ملاحظة: $e_i = y_i - \hat{y}_i$
2. ربّع كل خطأ: e_i^2
3. اجمع كل المربعات
4. اقسم على عدد الملاحظات

خصائص MSE:

- دائماً قيمة موجبة أو صفر
- الوحدة هي مربع وحدة المتغير التابع (مثلاً ريال² إذا كان المتغير بالريال)
- حساس جداً للقيم الشاذة (لأن التربيع يضخم الأخطاء الكبيرة)
- له خصائص رياضية مرغوبة (قابل للاشتقاق، محدب)
- أساسي في نظرية الانحدار الخطي

لماذا نربّع؟

التربيع يعطي عقوبة أكبر بكثير للأخطاء الكبيرة. خطأ بمقدار 10 يساهم بـ 100، بينما خطأ بمقدار 2 يساهم بـ 4 فقط. هذا مفيد عندما نريد تجنب الأخطاء الكبيرة بشدة.

Restricted - مقيد

عنوان | مثال عملي مقارن مع - MSEو MAE

نفس مثال التنبؤ بدرجات الحرارة:

اليوم	y_i	$\hat{y}_i$	e_i	$\lvert e_i \rvert$	e_i^2
1	25	23	2	2	4
2	28	30	-2	2	4
3	22	24	-2	2	4
4	30	28	2	2	4
5	26	26	0	0	0

حساب) MAEللمراجعة:

MAE = (2 + 2 + 2 + 2 + 0) / 5 = 8 / 5 = 1.6°C

حساب :MSE

الخطوة 1: جمع المربعات:

$\Sigma e_i^2 = 4 + 4 + 4 + 4 + 0 = 16$

الخطوة 2: القسمة على عدد الملاحظات:

MSE = 16 / 5 = 3.2 (°C)²

ملاحظة مهمة:

MSE = 3.2(درجة مئوية)² - الوحدة مربّعة!

مقارنة بسيطة:

- MAE = 1.6°C (سهل الفهم مباشرة)
- MSE = 3.2 (°C)² (أقل بداهة بسبب الوحدة المربعة)

الآن لنرى تأثير قيمة شاذة:

لنفترض أن اليوم السادس كان الفعلي 25° Cوالمتوقع 35°) Cخطأ = -10):

مع القيمة الشاذة:

- MAEالجديد = (2+2+2+2+0+10) / 6 = 18/6 = 3.0°) Cارتفع من 1.6 إلى 3.0)
- MSEالجديد = (4+4+4+4+0+100) / 6 = 116/6 = 19.33 (°)² (Cارتفع من 3.2 إلى 19.33)

الملاحظة الأساسية:

MSEتأثر بالقيمة الشاذة أكثر بكثير من) MAEزيادة بنسبة 504% مقابل 87.5%)!

Restricted - مقيد

عنوان | MSE - المزايا والعيوب

مزايا MSE:

1. يعاقب الأخطاء الكبيرة بشدة:

هذه الخاصية مفيدة جداً عندما تكون الأخطاء الكبيرة غير مقبولة أو خطيرة. في التطبيقات الحساسة (الطب، المالية، الأمان)، نريد تجنب الأخطاء الكبيرة حتى لو كان ذلك على حساب زيادة طفيفة في الأخطاء الصغيرة.

2. خصائص رياضية ممتازة:

- قابل للاشتقاق في كل نقطة (differentiable everywhere)
- دالة محدبة (convex)، مما يسهل إيجاد الحد الأدنى
- هذا يجعله مثالياً لخوارزميات التحسين في تدريب النماذج

3. الأساس النظري للانحدار الخطي:

طريقة المربعات الصغرى (Ordinary Least Squares - OLS) تعتمد على تقليل MSE. له أساس إحصائي عميق في نظرية التقدير.

4. تفكيك التباين:

MSE يرتبط مباشرة بمفاهيم التباين، مما يسمح بتحليلات إحصائية عميقة.

عيوب MSE:

1. الوحدة المربعة:

MSE بوحدة مربع المتغير الأصلي، مما يجعل التفسير المباشر صعباً. ماذا تعني "100 ريال^2"؟

2. حساسية شديدة للقيم الشاذة:

قيمة شاذة واحدة يمكن أن تضخم MSE بشكل كبير جداً، مما قد يعطي انطباعاً مضللاً عن الأداء العام للنموذج.

3. عدم التناسب في العقوبة:

في بعض الحالات، قد لا نريد عقوبة تربيعية. قد نفضل عقوبة أكثر توازناً أو حتى أقل للأخطاء الكبيرة (حسب السياق).

4. صعوبة المقارنة بين مشاكل مختلفة:

لأن الوحدة مربعة، من الصعب مقارنة MSE لنماذج تتنبأ بمتغيرات مختلفة.

Restricted - مقيد

عنوان | معامل التحديد R^2 - المقدمة والمفهوم

Coefficient of Determination (R^2):

السؤال الأساسي:

المعايير السابقة MAE, MSE تخبرنا "بكم" يخطئ النموذج، لكن لا تخبرنا "ما مدى جودته" مقارنة بخط الأساس. هنا يأتي دور R^2.

ما هو R^2؟

معامل التحديد R^2 هو معيار يقيس نسبة التباين في المتغير التابع التي يفسرها النموذج. بعبارة أخرى: "ما مقدار التغير في Y الذي يستطيع نموذجنا تفسيره؟"

التعريف الحدسي:

R^2 يجيب على السؤال: "ما مدى تحسن نموذجنا مقارنة بأبسط نموذج ممكن (استخدام المتوسط فقط)؟"

النطاق:

- R^2 يتراوح عادة بين 0 و 1
- $R^2 = 1$: نموذج مثالي (يفسر كل التباين)
- $R^2 = 0$: النموذج لا يفسر أي شيء (مثل استخدام المتوسط)
- $R^2 < 0$: النموذج أسوأ من استخدام المتوسط! (ممكن في بيانات الاختبار)

التفسير بالنسبة المئوية:

$R^2 = 0.85$ يعني أن النموذج يفسر 85% من التباين في البيانات، و15% المتبقية تعود لعوامل غير مدرجة في النموذج أو ضوضاء عشوائية.

أهمية R^2:

R^2 معيار مستقل نسبياً يمكن مقارنته عبر مشاكل مختلفة، على عكس MAE/MSE التي تعتمد على وحدة القياس وحجم البيانات.

Restricted - مقيد

عنوان | R^2 - الصيغة الرياضية والمكونات

الصيغة الرياضية:

$R^2 = 1 - (SS_res / SS_tot)$

حيث:

- **SS_res**: مجموع مربعات البواقي (Residual Sum of Squares)
- **SS_tot**: مجموع المربعات الكلي (Total Sum of Squares)

تفصيل المكونات:

1. SS_res (مجموع مربعات البواقي):

$SS_res = \Sigma(y_i - \hat{y}_i)^2$

هذا هو مجموع مربعات الأخطاء بين القيم الفعلية والمتوقعة. يقيس "الخطأ المتبقي" بعد تطبيق النموذج.

- ملاحظة: $SS_res = n \times MSE$

2. SS_tot (مجموع المربعات الكلي):

$SS_tot = \Sigma(y_i - \bar{y})^2$

حيث $\bar{y}$ هو متوسط القيم الفعلية. هذا يقيس التباين الكلي في البيانات - كم تتباعد القيم عن متوسطها.

3. SS_exp (مجموع المربعات المفسَّر):

$SS_exp = \Sigma(\hat{y}_i - \bar{y})^2$

يقيس التباين الذي "يفسره" النموذج.

العلاقة الأساسية:

$SS_tot = SS_exp + SS_res$

إعادة كتابة R^2:

$R^2 = SS_exp / SS_tot = 1 - (SS_res / SS_tot)$

التفسير:

- البسط (SS_res): ما لم يستطع النموذج تفسيره
- المقام (SS_tot): كل التباين الموجود
- النسبة (SS_res/SS_tot): نسبة التباين غير المفسَّر
- $R^2 = 1 -$ هذه النسبة = نسبة التباين المفسَّر

Restricted - مقيد

عنوان | R^2 - مثال تفصيلي خطوة بخطوة

مثال: التنبؤ بأسعار منازل (بالألف ريال)

المنزل	السعر الفعلي (y_i)	السعر المتوقع($\hat{y}_i$)
1	200	195
2	250	240
3	300	310
4	350	345
5	400	410

الخطوة 1: حساب المتوسط $\bar{y}$

$\bar{y} = (200 + 250 + 300 + 350 + 400) / 5 = 1500 / 5 = 300$

الخطوة 2: حساب SS_tot

y_i	$(y_i - \bar{y})$	$(y_i - \bar{y})^2$
200	-100	10,000
250	-50	2,500
300	0	0
350	50	2,500
400	100	10,000

$SS_tot = 10,000 + 2,500 + 0 + 2,500 + 10,000 = 25,000$

- Restricted مقيد

عنوان | مثال تفصيلي خطوة بخطوة - R^2

الخطوة 3: حساب SS_res

y_i	$\hat{y}_i$	$(y_i - \hat{y}_i)$	$(y_i - \hat{y}_i)^2$
200	195	5	25
250	240	10	100
300	310	-10	100
350	345	5	25
400	410	-10	100

SS_res = 25 + 100 + 100 + 25 + 100 = 350

الخطوة 4: حساب R^2

$R^2 = 1 - (SS_res / SS_tot)$

$R^2 = 1 - (350 / 25{,}000)$

$R^2 = 1 - 0.014$

$R^2 = 0.986$

التفسير:

$R^2 = 0.986$ أو 98.6%

النموذج يفسر 98.6% من التباين في أسعار المنازل. هذا نموذج ممتاز جداً! فقط 1.4% من التباين غير مفسَّر.

Restricted - مقيد

عنوان | R^2 التفسير والفهم العميق

تفسير قيم R^2:

R^2 = 1.0 (أو 100%):

- نموذج مثالي تماماً
- كل نقطة على خط الانحدار بالضبط
- SS_res = 0
- نادر جداً في الواقع (قد يشير لـ overfitting)

R^2 = 0.9 إلى 0.99 (90%-99%):

- نموذج ممتاز جداً
- يفسر معظم التباين
- مناسب جداً للاستخدام العملي
- شائع في البيانات الفيزيائية أو المخبرية المنضبطة

R^2 = 0.7 إلى 0.9 (70%-90%):

- نموذج جيد إلى جيد جداً
- يفسر جزءاً كبيراً من التباين
- مقبول في كثير من التطبيقات
- شائع في العلوم الاجتماعية والاقتصاد

R^2 = 0.5 إلى 0.7 (50%-70%):

- نموذج متوسط
- يفسر نصف التباين أو أكثر
- قد يكون مفيداً حسب السياق
- يشير إلى وجود عوامل مهمة غير مدرجة

R^2 = 0.3 إلى 0.5 (30%-50%):

- نموذج ضعيف
- لكن قد يكون مفيداً في بعض المجالات المعقدة
- يحتاج لتحسين (ميزات إضافية، نموذج أفضل)

R^2 < 0.3 (أقل من 30%):

- نموذج ضعيف جداً
- تقريباً لا قيمة عملية له
- يحتاج لإعادة نظر كاملة

R^2 = 0:

- النموذج مثل استخدام المتوسط فقط
- لا يفسر أي تباين
- عديم الفائدة

R^2 سالب (R^2 < 0):

- النموذج أسوأ من استخدام المتوسط!
- يحدث عادة في بيانات الاختبار لنموذج سيئ جداً
- علامة خطر كبيرة - النموذج فاشل

ملاحظة مهمة جداً:

قيمة " R^2 الجيدة" تعتمد على المجال:

- الفيزياء/الهندسة: قد نتوقع R^2 > 0.95
- الطب/البيولوجيا: R^2 > 0.7 قد يكون ممتازاً
- العلوم الاجتماعية: R^2 > 0.5 قد يكون جيداً
- السلوك البشري: R^2 > 0.3 قد يكون مقبولاً

Restricted - مقيد

عنوان | R^2 - المزايا والعيوب والمحاذير

مزايا R^2:

1. معيار نسبي موحد:

بعكس R^2، MAE/MSE/ لا يعتمد على وحدة القياس. يمكن مقارنة R^2 لنموذج يتنبأ بالأسعار مع R^2 لنموذج يتنبأ بالأوزان.

2. تفسير حدسي:

"النموذج يفسر X% من التباين" - سهل الفهم حتى لغير المتخصصين.

3. يقيس جودة النموذج نسبياً:

يقارن النموذج بخط الأساس (المتوسط)، مما يعطي فكرة عن "القيمة المضافة" للنموذج.

4. تحليل المساهمات:

في الانحدار المتعدد، يمكن استخدام R^2 لتحليل مساهمة المتغيرات المختلفة.

عيوب R^2 ومحاذير مهمة:

1. يزداد دائماً بإضافة متغيرات:

إضافة أي متغير (حتى لو عشوائي!) لن يُنقص R^2 أبداً، بل قد يزيده قليلاً. هذا مضلل!

الحل: استخدام R^2 المعدَّل (Adjusted R^2) الذي يعاقب على إضافة متغيرات غير مفيدة.

2. لا يكشف الإفراط في التعلم مباشرة:

R^2 عالي على البيانات التدريبية لا يضمن أداءً جيداً على بيانات جديدة.

الحل: احسب R^2 على بيانات الاختبار أيضاً.

3. لا يكشف العلاقات غير الخطية:

إذا كانت العلاقة غير خطية، R^2 المحسوب للنموذج الخطي قد يكون منخفضاً حتى لو كانت العلاقة قوية.

4. حساس للقيم الشاذة:

مثل R^2، MSE يعتمد على المربعات، فهو حساس للقيم الشاذة.

5. لا يكشف التحيز:

R^2 عالي لا يعني أن النموذج غير متحيز. قد يكون النموذج يتنبأ بدقة نسبية لكن دائماً يقلل أو يزيد التقدير.

6. السياق ضروري:

$R^2 = 0.6$ قد يكون ممتازاً في مجال وضعيفاً في آخر. لا توجد "قيمة سحرية" عامة.

7. لا يحل محل الفحص البصري:

دائماً ارسم البيانات والتنبؤات! R^2 عالي لا يعني أن النموذج صحيح أو مناسب.

Restricted - مقيد

عنوان | المراجع

Practical Statistics for Data Scientists – Peter Bruce

Restricted - مقيد

لكم شكراً