

المحور الأول: أساسيات التعلم الآلي

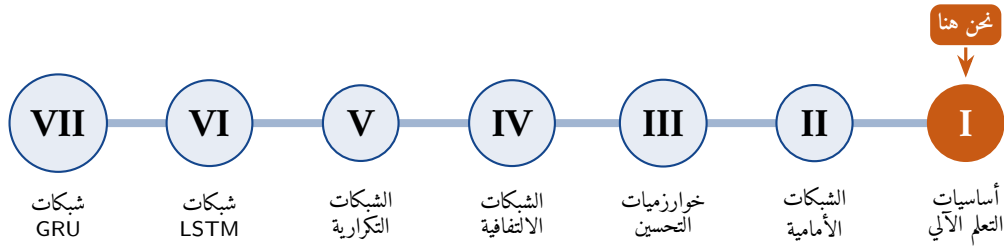
نماذج التصنيف

مقياس: تطبيقات التعلم العميق
الأستاذ(ة): د. غيلوس عز الدين

ماستر 2: مالية وتجارة دولية -- تسويق
كلية العلوم الاقتصادية والتجارية وعلوم التسيير
جامعة عبد الحفيظ بوالصوف -- ميله

السنة الجامعية 2026/2027

خريطة المقياس: أين نحن الآن؟



- التعلم الموجه ينقسم إلى تصنيف (مخرج فقوي) وانحدار (مخرج عددي).
- التصنيف يُستعمل حين يكون السؤال من نوع «هل...؟» بعدد محدود من الإجابات.

اليوم

سندخل بالتفصيل إلى عالم التصنيف: نماذج، وكيف نعرف أن نموذج التصنيف جيد فعلاً.

- 1 التصنيف الثنائي والمتعدد: تعريف موسّع وأمثلة.
- 2 حدّ الفصل (Decision Boundary).
- 3 أهم نماذج التصنيف: الانحدار اللوجستي، k-NN، أشجار القرار، الشبكات العصبية.
- 4 مصفوفة الالتباس ومكوناتها.
- 5 مقاييس التقييم: الدقة، Precision، Recall، F1.
- 6 لماذا لا تكفي الدقة؟ واختيار المقياس المناسب.

في نهاية هذه المحاضرة ستكون قادراً على:

- التمييز بين التصنيف الثنائي والمتعدد الفئات.
- شرح الفكرة الحدسية لأربعة نماذج تصنيف شائعة.
- قراءة مصفوفة الالتباس وحساب مقاييس الأداء يدوياً.
- اختيار المقياس المناسب حسب طبيعة المسألة وكلفة الخطأ.

التعريف

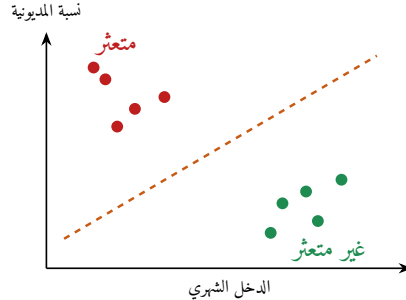
مهمة تهدف إلى تصنيف كل مثال ضمن فئة من فئات محددة مسبقاً انطلاقاً من خصائصه (المدخلات).

- تصنيف ثنائي: فئتان فقط (تعثر / عدم تعثر).
- تصنيف متعدد الفئات: أكثر من فئتين (نوع منتج، لغة نص، شريحة زبون).
- تصنيف متعدد التسميات: مثال واحد قد ينتمي لعدة فئات معاً.

أمثلة: ثنائي مقابل متعدد الفئات

المسألة	النوع	الفئات الممكنة
هل سيشتري الزبون المنتج؟	ثنائي	نعم / لا
هل البريد مزعج؟	ثنائي	مزعج / عادي
تصنيف فئة المنتج	متعدد الفئات	إلكترونيات / ملابس / أخرى
تصنيف مشاعر تعليق	متعدد الفئات	إيجابي / سلبي / محايد
تصنيف السلعة جمركياً	متعدد الفئات	عشرات الرموز الجمركية

حدّ الفصل (Decision Boundary)

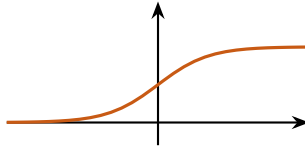


الفكرة: يبحث النموذج عن حدّ فاصل (خط، منحنى، أو سطح) يفصل الفئتين بأقل خطأ ممكن.

النموذج 1: الانحدار اللوجستي

الفكرة

يحسب احتمال انتماء المثال لفئة معينة باستعمال دالة Sigmoid، ثم يقارنه بعتبة (غالباً 0.5).



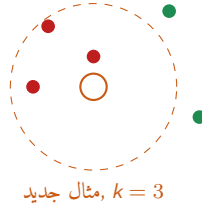
مثال تطبيقي

احتمال تعثر 0.78 $\leftarrow$ أعلى من العتبة 0.5 $\leftarrow$ يُصنّف «متعثر».

النموذج 2: أقرب الجيران (k-NN)

الفكرة

يصنّف المثال الجديد حسب الفئة الأغلب بين أقرب k أمثلة له في بيانات التدريب.

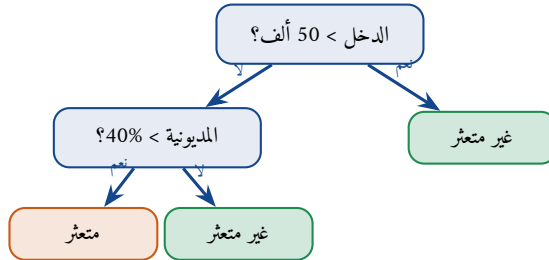


مثال تطبيقي

تصنيف زبون جديد حسب سلوك أقرب 5 زبائن له في الدخل والعمر وحجم المشتريات.

الفكرة

سلسلة من الأسئلة «نعم/لا» على الخصائص تؤدي تدريجياً إلى القرار النهائي.



الفكرة

طبقات من العصبونات تتعلم تدريجياً حدود فصل معقدة وغير خطية، وتُخرج احتمال الفئة عبر دالة Sigmoid (ثنائي) أو Softmax (متعدد الفئات).

- تتفوق خصوصاً عندما تكون البيانات كبيرة وغير منظمة (نصوص، صور).
- سنفصلها في المحاور القادمة (الشبكات الأمامية وما بعدها).

مقارنة نماذج التصنيف الأربعة

النموذج	الفكرة	قابلية التفسير	حجم البيانات
الانحدار اللوجستي	احتمال خطي	عالية	صغير-متوسط
k-NN	أقرب الأمثلة	متوسطة	صغير
أشجار القرار	أسئلة متتالية	عالية جداً	متوسط
الشبكات العصبية	طبقات معقدة	منخفضة	كبير

من التصنيف إلى التقييم: لماذا؟

السؤال الأهم

بنينا نموذجاً، لكن كيف نعرف أنه جيد؟ الدقة البصرية على رسم بسيط لا تكفي؛ نحتاج أرقاماً دقيقة.

- الأداة الأساسية: مصفوفة الالتباس.
- منها نشق كل مقاييس التقييم الأخرى.

مصفوفة الالتباس (Confusion Matrix)

القيمة المتوقعة			
		سلي	إيجابي
القيمة الفعلية	إيجابي	سلي خاطئ FN	إيجابي صحيح TP
	سلي	سلي صحيح TN	إيجابي خاطئ FP

معنى كل خلية في سياق كشف الاحتيال

TP	معاملة احتيالية فعلاً، وتنبأ النموذج بأنها احتيالية (صحيح)
FP	معاملة سليمة، لكن تنبأ النموذج بأنها احتيالية (تنبيه كاذب)
FN	معاملة احتيالية فعلاً، لكن النموذج لم يكتشفها (خطر جداً)
TN	معاملة سليمة، وتنبأ النموذج بأنها سليمة (صحيح)

الصيغة

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

نسبة التنبؤات الصحيحة من مجموع التنبؤات.

تنبيه

مضللة في البيانات غير المتوازنة. مثال: من 1000 معاملة، 10 فقط احتيالية. نموذج «كسول» يصنف الجميع كـ«سليم» يحقق دقة 99% دون اكتشاف أي احتيال!

Precision (الدقة الموجبة)

$$\text{Precision} = \frac{TP}{TP + FP}$$

من بين ما توقعه النموذج كإيجابي، كم كان صحيحاً فعلاً؟

Recall (الاستدعاء/الحساسية)

$$\text{Recall} = \frac{TP}{TP + FN}$$

من بين الحالات الإيجابية الحقيقية، كم اكتشف النموذج؟

الصيغة

$$F1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

متوسط توافقي بين Precision و Recall؛ مفيد عندما نريد توازناً بينهما.

مثال تطبيقي

نموذج Precision=0.9 لكن Recall=0.2 يبدو جيداً بمقياس واحد، لكن F1 منخفض يكشف أن النموذج يفوت أغلب الحالات الإيجابية الحقيقية.

أي مقياس أهم حسب السياق؟

السياق	الأولوية
كشف الاحتيال (لا نريد تفويت حالة احتيال)	Recall عالٍ
الموافقة على قرض (لا نريد رفض عميل جيد خطأً)	Precision عالٍ
تشخيص طبي أولي	Recall عالٍ
توصية منتج (كلفة الخطأ منخفضة)	الدقة تكفي غالباً

المعطيات

اختبرنا نموذج كشف احتيال على 200 معاملة: 20 احتيالية فعلاً. النموذج اكتشف 15 حالة صحيحة (TP)، فاته 5 (FN)، وأخطأ في تصنيف 10 معاملات سليمة كاحتيال (FP).

- 1 احسب TN، ثم ارسم مصفوفة الالتباس كاملة.
- 2 احسب الدقة، Precision، Recall، وF1.
- 3 هل تنصح البنك باعتماد هذا النموذج؟ علّل.

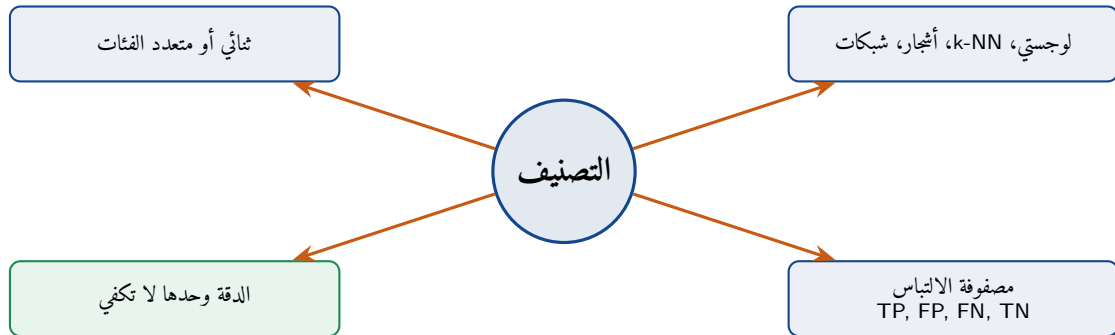
$$170 = 10 - 5 - 15 - 200 = \text{TN}$$

$$\text{Accuracy} = \frac{15 + 170}{200} = 0.925$$

$$\text{Precision} = \frac{15}{25} = 0.60 \quad \text{Recall} = \frac{15}{20} = 0.75$$

$$F1 = 2 \times \frac{0.60 \times 0.75}{0.60 + 0.75} \approx 0.67$$

- الدقة الظاهرية مرتفعة (92.5%) لكنها مضللة بسبب قلة حالات الاحتيال.
- $0.75 = \text{Recall}$ يعني أن 25% من حالات الاحتيال الفعلية ما زالت تفلت.



للمناقشة

- (1) ما الفرق بين Precision و Recall، ومتى نفضل أحدهما على الآخر؟
- (2) لماذا لا تكفي الدقة في حالة كشف الاحتيال؟
- (3) اقترح مثلاً من تخصصك يحتاج Recall عالياً، وآخر يحتاج Precision عالياً.

المراجع

- Ketkar, N. (2017). Deep Learning with Python. Apress. Chapter 2.
- Goodfellow, I., Bengio, Y., Courville, A. (2016). Deep Learning. MIT Press.

الأسبوع القادم

المحور الأول (ج): الانحدار وتقييم الأداء الشامل — $MAE/RMSE/R^2$ ، فرط التخصيص، وتقسيم البيانات.