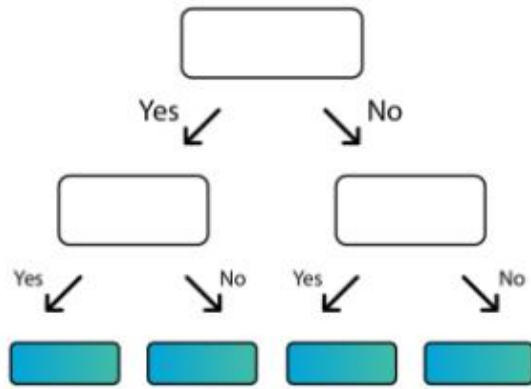


شجرة القرار (DECISION TREE)

خوارزمية شجرة القرار (Decision Tree Algorithm)

هي خوارزمية تعلم تحت الاشراف (Supervised Learning) ذات بنية شجرية تشبه بنية المخططات التدفقية (Flowcharts) وتستخدم في التصنيف (Classification) والانحدار (Regression) حيث يتم استخدامها لتصنيف العناصر إلى فئات مختلفة بناءً على متغيرات معينة. كما يمكن استخدامها في تحليل البيانات واكتشاف العلاقات السببية بين المتغيرات.



Decision Tree

شجرة القرار

الهدف العام لاستخدام خوارزمية شجرة القرار هو إنشاء نموذج تدريب يمكن استخدامه للتنبؤ بفئة أو قيمة المتغيرات المستهدفة من خلال تعلم قواعد القرار المستنتجة من البيانات السابقة (بيانات التدريب). حيث يتم اعتبار مجموعة البيانات التدريبية على أنها أسئلة تجاب بنعم أو لا وتستمر البيانات بالتفرع إلى أن يتم الوصول إلى النتيجة النهائية. شجرة القرار أيضا تعتبر أنها مجموعة عمليات من If else حيث أنه كل فرع من فروع الشجرة هو حالة شرطية وتكون النتيجة النهائية في أطراف الفروع.

شجرة القرار (DECISION TREE)

الأنواع الرئيسية لأشجار القرار (Types of Decision Trees)

1. شجرة التصنيف (Classification Tree)

وتتضمن متغيرات هدف مقسمة إلى فئات؛ أي أن المخرج أو القرار قطعي يعبر عن فئات محددة مثل نعم أو لا. وتستخدم مثلاً للتنبؤ بلياقة شخص ما بناءً على معلومات مثل العمر والعادات الغذائية والنشاط الجسدي، فالخرج هنا سيكون محدد إما لائق أو غير لائق.

2. شجرة الانحدار (Regression Tree)

وتتضمن متغيرات هدف مستمرة؛ أي أن المخرج أو القرار مُستمر مثل قيمة عددية ضمن مجال محدد. وتستخدم مثلاً للتنبؤ بدخل شخص ما بناءً على مجموعة من المتغيرات المستمرة مثل المنصب الذي يشغله وخبرته وما إلى ذلك.

شجرة القرار (DECISION TREE)

بنية شجرة القرار (Decision Tree Structure)

- **العقدة الجذرية (Root Node)**
نقطة البداية في الشجرة، وتمثل مجموعة البيانات بأكملها. عندها يتم اتخاذ القرار الأول. تحتوي العقدة الجذرية على السؤال أو الانقسام الأول بناءً على أهم خاصية في مجموعة البيانات.
- **الفروع (Branches)**
تمثل النتائج المحتملة لقرار أو قاعدة اختبار، وتقود إلى عقدة قرار أخرى أو إلى نتيجة نهائية. كل فرع يمثل مسار قرار يؤدي إما إلى عقدة قرار أخرى أو إلى عقدة ورقة.
- **العقد الداخلية / عقد القرار (Internal / Child / Decision Nodes)**
نقاط يتم فيها تقسيم مجموعة البيانات بناءً على سمات محددة، مما يؤدي إلى المزيد من الفروع. تحتوي هذه العقد على أسئلة أو معايير لتقسيم البيانات إلى مجموعات فرعية مختلفة.
- **عقد الأوراق / العقد النهائية (Leaf Nodes / Terminal Nodes)**
النتائج النهائية لمسار اتخاذ القرار، وتمثل تصنيفًا أو قرارًا. بمجرد أن يصل مسار القرار إلى عقدة ورقية، يتم إصدار التنبؤ.

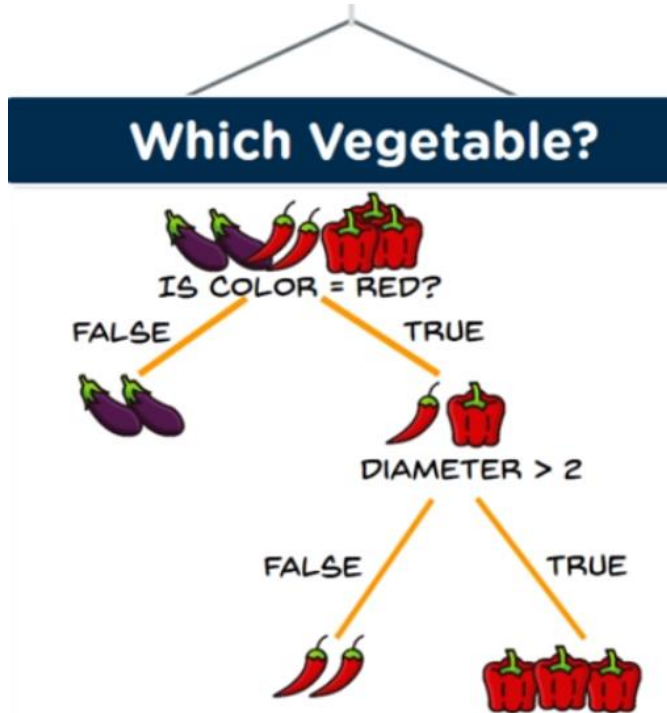
شجرة القرار (DECISION TREE)

المبدأ العام لشجرة القرار (Decision Tree)

- تعمل شجرة القرار على تمثيل عملية اتخاذ القرار بشكل هرمي شجري؛ حيث تبدأ البيانات في عقدة الجذر (Root Node) على شكل مجموعة كاملة، ثم تُقسَّم هذه البيانات تدريجيًا إلى مجموعات أصغر في عُقد فرعية (Child Nodes) اعتمادًا على قيمة سمة معيّنة (Feature/Attribute) في كل خطوة.
- في كل عقدة قرار (Decision Node) يتم اختيار سمة للتقسيم، فيُقسَّم المسار إلى فروع (Branches) وفقًا لقيم هذه السمة، ويستمر هذا الانقسام المتتابع حتى نصل إلى العُقد الورقية (Leaf Nodes)، وهي العقد النهائية التي تُمثِّل النتيجة أو التصنيف (Class/Label) الذي تُرجعه الشجرة.

شجرة القرار للتصنيف (Classification Tree)

مثال على شجرة قرار للتصنيف (Classification Tree):



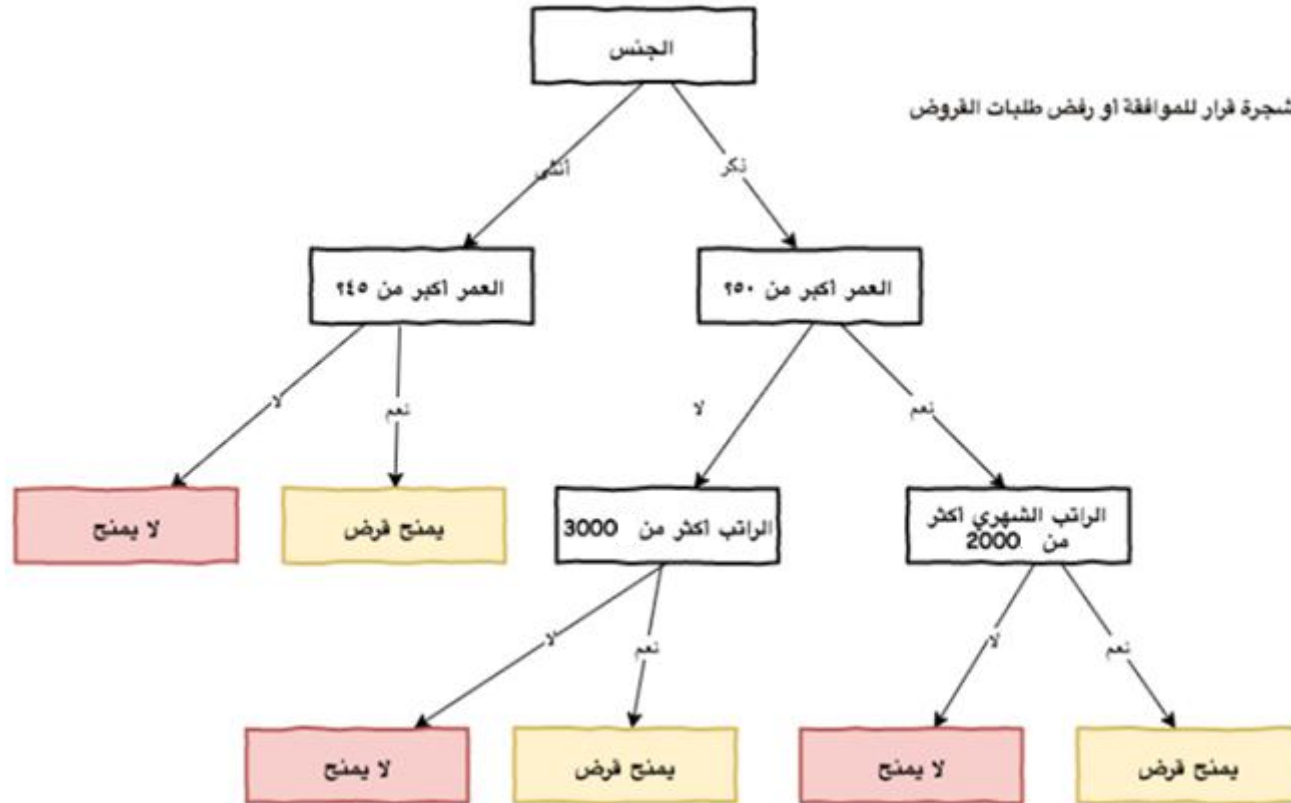
تعمل خوارزمية شجرة القرار عن طريق تحليل البيانات واستخراج العلاقات (السمات) بين المتغيرات المختلفة للوصول إلى قرار تصنيف دقيق.

بفرض أننا نريد تصنيف مجموعة من الخضار بالاعتماد على لون وقطر الخضار كما في الصورة التالية :

وكما نلاحظ تم التصنيف بشرط أن يكون لون الخضار أحمر وقطره يتجاوز ال 2 ، ولكن كيف نضع هذه الشروط ؟ كيف نختارها بحيث نحصل على أفضل دقة ؟

شجرة القرار للتصنيف (Classification Tree)

مثال على شجرة قرار للتصنيف (Classification Tree):



لبناء النموذج الخاص باعطاء القروض نحتاج إلى تقسيم العملاء إلى فئات بحسب الجنس والدخل لأجل اتخاذ القرار في منح القرض من عدمه. يتم تمثيل الشجرة كما هو موضح في الشكل التالي:

كما هو واضح في المثال، ربما يكون واحد من أهم التحديات في هذا الأسلوب هو اختيار الأسئلة أو النقاط التي من خلالها نستطيع تقسيم السجلات إلى فروع الشجرة ومن ثم الوصول للقرار.

شجرة القرار للتصنيف (Classification Tree)

أين تُستخدم أشجار القرار في الحياة الحقيقية؟

- تصنيف الرسائل: Spam / Not Spam
- في الطب: مساعدة الطبيب يقرّر إن كان المريض عالي خطورة أو منخفض بناءً على مجموعة فحوصات
- في التسويق: هل نعرض لهذا المستخدم عرضاً ترويجياً معيناً؟ بناءً على العمر، البلد، سلوك الشراء
- في البنوك: شجرة قرار لتقدير إذا كان الشخص مؤهلاً للحصول على قرض أو لا.

شجرة القرار = طريقة بسيطة ومرنة لتحويل البيانات إلى قرارات حقيقية.

شجرة القرار للتصنيف (Classification Tree)

مثال:

ابن شجرة قرار لاختيار التخصص الجامعي المناسب

وصف المهمة:

باستخدام مبدأ أشجار القرار (Decision Trees)، قم بتنفيذ ما يلي:

1. تخيل أنك تساعد طالبًا في اختيار تخصصه الجامعي.
2. اعتمد في أسئلتك على مجموعة من المعايير، مثل:

- حب الرياضيات
- حب البرمجة والعمل على الحاسوب
- حب الرسم والتصميم
- حب النقاش والتعامل مع الناس

شجرة القرار للتصنيف (Classification Tree)

3. ارسم شجرة قرار على شكل مخطط هرمي يبدأ بسؤال عام في الأعلى مثل:

"هل تحب الرياضيات؟"

4. من كل سؤال، ارسم أسهمًا تمثل الإجابات الممكنة (نعم / لا)، وانتقل إلى أسئلة أخرى أكثر تخصصًا.

5. في نهاية كل فرع من فروع الشجرة، اكتب اسم تخصص مقترح مناسب، مثل:

- هندسة حاسوب (Computer Engineering)
- ذكاء اصطناعي (Artificial Intelligence)
- تصميم جرافيك (Graphic Design)
- قانون (Law)

شجرة القرار للتصنيف (Classification Tree)

6. تأكد من:

- وضوح الأسئلة وترتيبها المنطقي.
- وضوح المسارات من الجذر حتى الأوراق (Leaves)
- أن كل ورقة تمثل قرارًا نهائيًا: تخصص جامعي مقترح.

شجرة القرار للتصنيف (Classification Tree)

الحل:

1. تحديد المعايير (Features)

سنستخدم ٣ معايير بسيطة مرتبطة باختيار التخصص:

- حب الرياضيات

نعم / لا

- حب البرمجة والعمل على الحاسوب لفترات طويلة

نعم / لا

- حب الرسم والتصميم الفني

نعم / لا

- (احتياطي في فرع الآداب): حب النقاش والقضايا القانونية والتعامل مع الناس

نعم / لا

لن نستخدمها كلها في مستوى واحد، بل سنوزعها على شكل أسئلة متتابعة كما في شجرة القرار.

شجرة القرار للتصنيف (Classification Tree)

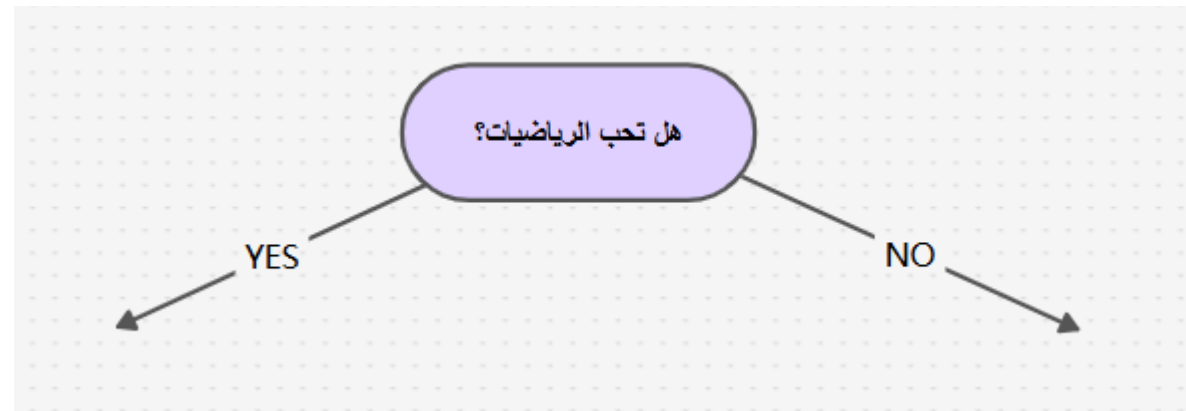
2. بناء شجرة القرار خطوة خطوة

الجذر (Root Node)

السؤال 1 (الجذر):

هل تحب الرياضيات؟

- إن كانت الإجابة: نعم --> نذهب إلى سؤال عن البرمجة
- إن كانت الإجابة: لا --> نذهب إلى سؤال عن الرسم والتعامل مع الناس



شجرة القرار للتصنيف (Classification Tree)

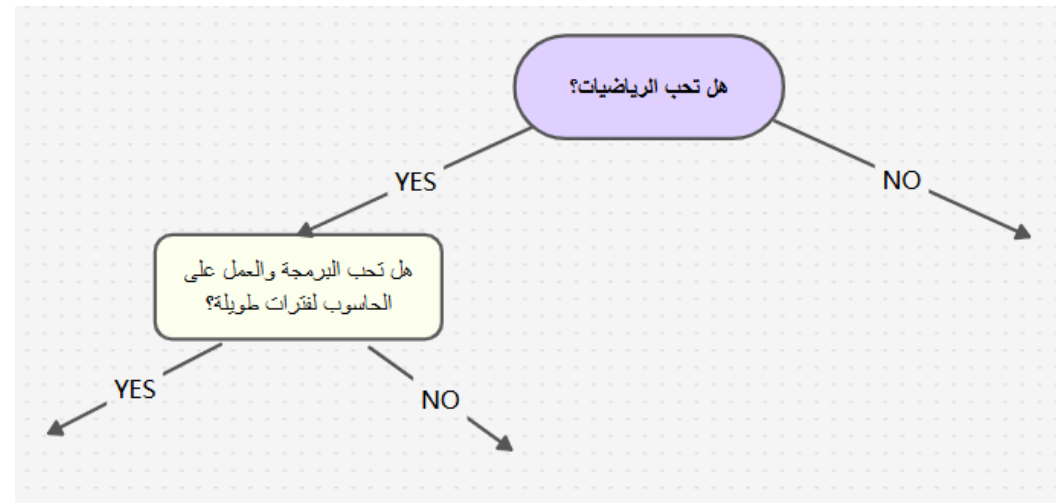
2. بناء شجرة القرار خطوة خطوة

الفرع الأيسر: يحب الرياضيات

السؤال 2:

هل تحب البرمجة والعمل على الحاسوب لفترات طويلة؟

- نعم --> نميّز بين هندسة حاسوب وذكاء اصطناعي
- لا --> نبحث عن تخصص أقرب لمهارات أخرى (هنا ممكن نذكر هندسة أخرى أو اقتصاد/محاسبة، لكن حسب التخصصات التي اتفقنا عليها سنركّز على التخصصات المقترحة)



شجرة القرار للتصنيف (Classification Tree)

إذا كانت الإجابة: نعم (يحب البرمجة)

السؤال 3:

هل تميل أكثر إلى الهاردوير والأنظمة والبنية التحتية، أم إلى الذكاء الاصطناعي والبيانات؟

- أميل إلى الأنظمة والهاردوير والشبكات

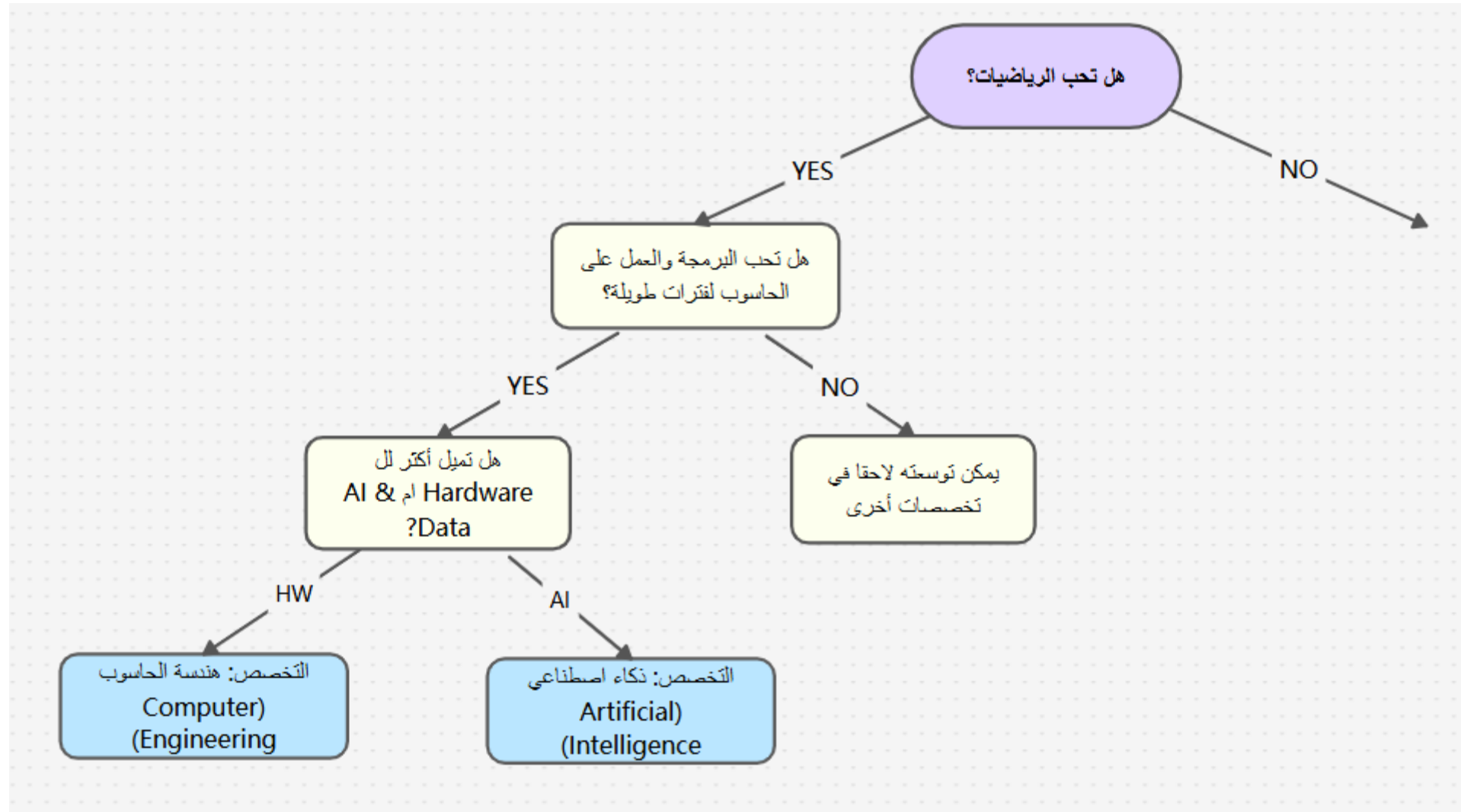
✓ تخصص مقترح: هندسة حاسوب (Computer Engineering)

- أميل إلى الذكاء الاصطناعي والبيانات والتعلم الآلي

✓ تخصص مقترح: ذكاء اصطناعي (Artificial Intelligence)

إذا كانت الإجابة: لا (يحب الرياضيات لكن لا يحب البرمجة كثيرًا)

- يمكن أن تُوسَّع لاحقًا لتخصصات مثل: رياضيات / إحصاء / علوم بيانات غير برمجية



شجرة القرار للتصنيف (Classification Tree)

الفرع الأيمن: لا يحب الرياضيات

السؤال 2 (في هذا الفرع):

هل تحب الرسم والتصميم الفني؟

- نعم --> نرشحه لتصميم جرافيك
 - لا --> نذهب لسؤال عن التعامل مع الناس والقضايا
- إذا كانت الإجابة: نعم (يحب الرسم)

✓ تخصص مقترح: تصميم جرافيك (Graphic Design)

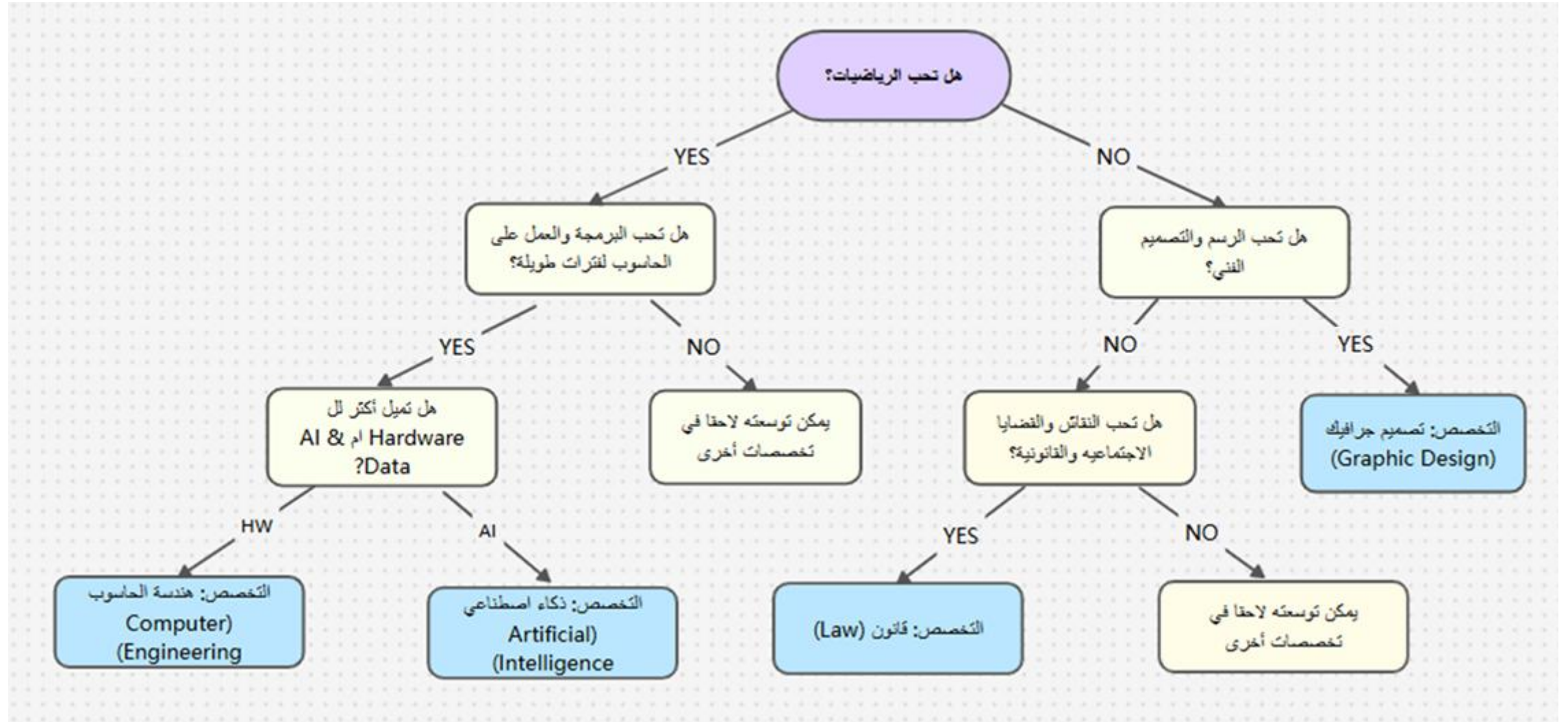
إذا كانت الإجابة: لا (لا يحب الرسم كثيرًا)

السؤال 3:

هل تحب النقاش، والحديث مع الناس، والقضايا الاجتماعية والقانونية، وتحب الإقناع والحجة؟

- نعم --> تخصص مقترح: قانون (Law)
- لا --> يمكن توسعة الشجرة لتخصصات أخرى (إدارة، تربية، لغات...)

شجرة القرار للتصنيف (Classification Tree)



شجرة القرار للتصنيف (Classification Tree)

مثال: التنبؤ بشراء منتج باستخدام شجرة القرار

ارسم شجرة قرار تتنبأ هل العميل سيشتري المنتج أو لن يشتري اعتمادًا على: (الدخل، العمر، وعدد المشتريات السابقة) بناءً على Dataset التالية

رقم العميل	الدخل (USD)	العمر (سنة)	عدد المشتريات السابقة	قرار الشراء (الهدف)
C1	40000	25	0	لا يشتري
C2	52000	28	1	لا يشتري
C3	70000	35	0	لا يشتري
C4	80000	40	3	يشتري
C5	30000	45	2	لا يشتري
C6	60000	32	5	يشتري
C7	48000	29	2	لا يشتري
C8	90000	38	0	لا يشتري

شجرة القرار للتصنيف (Classification Tree)

الحل:

تجهيز السمات (Features) للقرار

حتى نبني الشجرة بشكل بسيط، نحول القيم إلى أسئلة "نعم / لا":

- السمة 1: الدخل

$(\text{الدخل} < 50000) = \text{HighIncome}$

- السمة 2: العمر

$(\text{العمر} < 30) = \text{AgeAbove30}$

- السمة 3: المشتريات السابقة

$(\text{عدد المشتريات السابقة} > 0) = \text{HasPrevPurchase}$

شجرة القرار للتصنيف (Classification Tree)

جدول مختصر لهذه السمات:

العميل	HighIncome (الدخل < \$50000)	AgeAbove30 (العمر < 30)	HasPrevPurchase (<0؟)	قرار الشراء
C1	لا	لا	لا	لا يشتري
C2	نعم	لا	نعم	لا يشتري
C3	نعم	نعم	لا	لا يشتري
C4	نعم	نعم	نعم	يشترى
C5	لا	نعم	نعم	لا يشتري
C6	نعم	نعم	نعم	يشترى
C7	لا	لا	نعم	لا يشتري
C8	نعم	نعم	لا	لا يشتري

شجرة القرار للتصنيف (Classification Tree)

الخطوة 1 - اختيار العقدة الجذرية (Root)

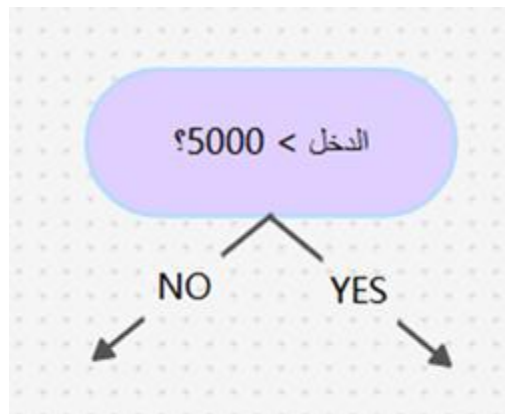
ننظر إلى السمة التي تقسم البيانات بشكل أوضح:

- عندما يكون الدخل ≥ 50000 :

--> C1, C5, C7 كلهم لا يشتري.

- عندما يكون الدخل < 50000 :

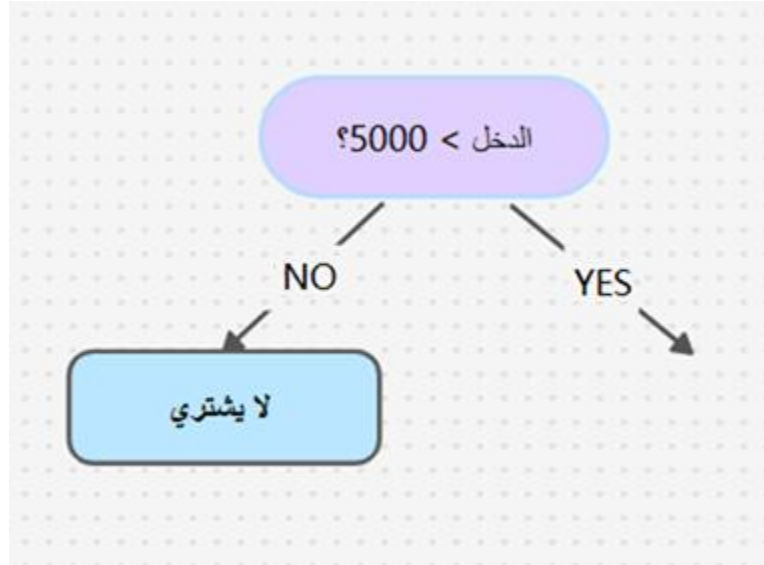
--> C2, C3, C4, C6, C8 هنا النتائج مختلطة (بعضهم يشتري، وبعضهم لا).



إذن سؤال الجذر المناسب:

هل الدخل أكبر من 50000؟

شجرة القرار للتصنيف (Classification Tree)



- الفرع لا "NO" يعطي مباشرة مجموعة كلها "لا يشتري" (ورقة نقية).
- الفرع نعم "YES" يحتاج أسئلة إضافية لتوضيح القرار.

هذا بالضبط ما تفعله خوارزميات أشجار القرار: تختار السمة التي تجعل المجموعات أكثر تجانساً في الفئة (نعم/لا).

شجرة القرار للتصنيف (Classification Tree)

الخطوة 2 – العقدة التالية (Age > 30) ؟

نركّز الآن على العملاء الذين دخلهم < 50000 :

العميل	HighIncome	AgeAbove30	HasPrevPurchase	قرار الشراء
C2	نعم	لا	نعم	لا يشتري
C3	نعم	نعم	لا	لا يشتري
C4	نعم	نعم	نعم	يشترى
C6	نعم	نعم	نعم	يشترى
C8	نعم	نعم	لا	لا يشتري

شجرة القرار للتصنيف (Classification Tree)

نلاحظ:

- عندما يكون **AgeAbove30 = لا** (العمر ≥ 30):

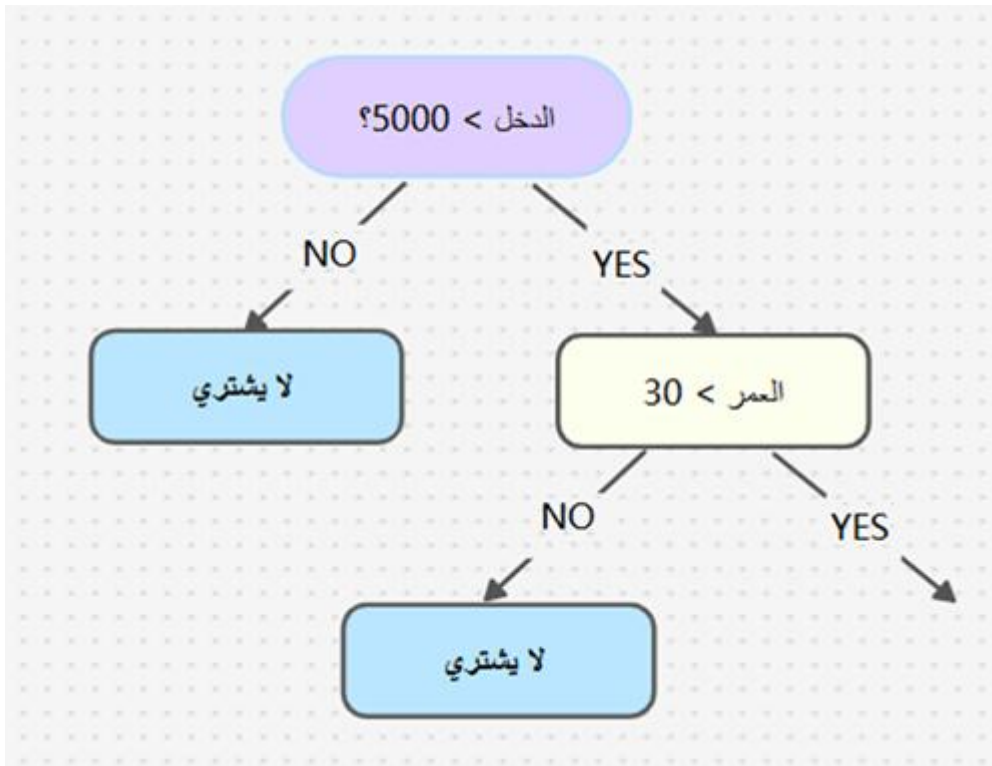
فقط C2، وقراره لا يشتري --> مجموعة نقية.

- عندما يكون **AgeAbove30 = نعم** (العمر < 30):

عندنا C3, C4, C6, C8، والنتائج مختلطة (يشتري / لا يشتري).

إذن السؤال في العقدة الثانية:

هل العمر أكبر من 30؟



شجرة القرار للتصنيف (Classification Tree)

الخطوة 3 – العقدة الثالثة (HasPrevPurchase)؟

نركّز على الجزء الذي فيه:

- الدخل $50000 <$

- العمر $30 <$

أي العملاء: C3, C4, C6, C8

العميل	HasPrevPurchase (<0)	قرار الشراء
C3	لا	لا يشتري
C4	نعم	يشترى
C6	نعم	يشترى
C8	لا	لا يشتري

شجرة القرار للتصنيف (Classification Tree)

نلاحظ:

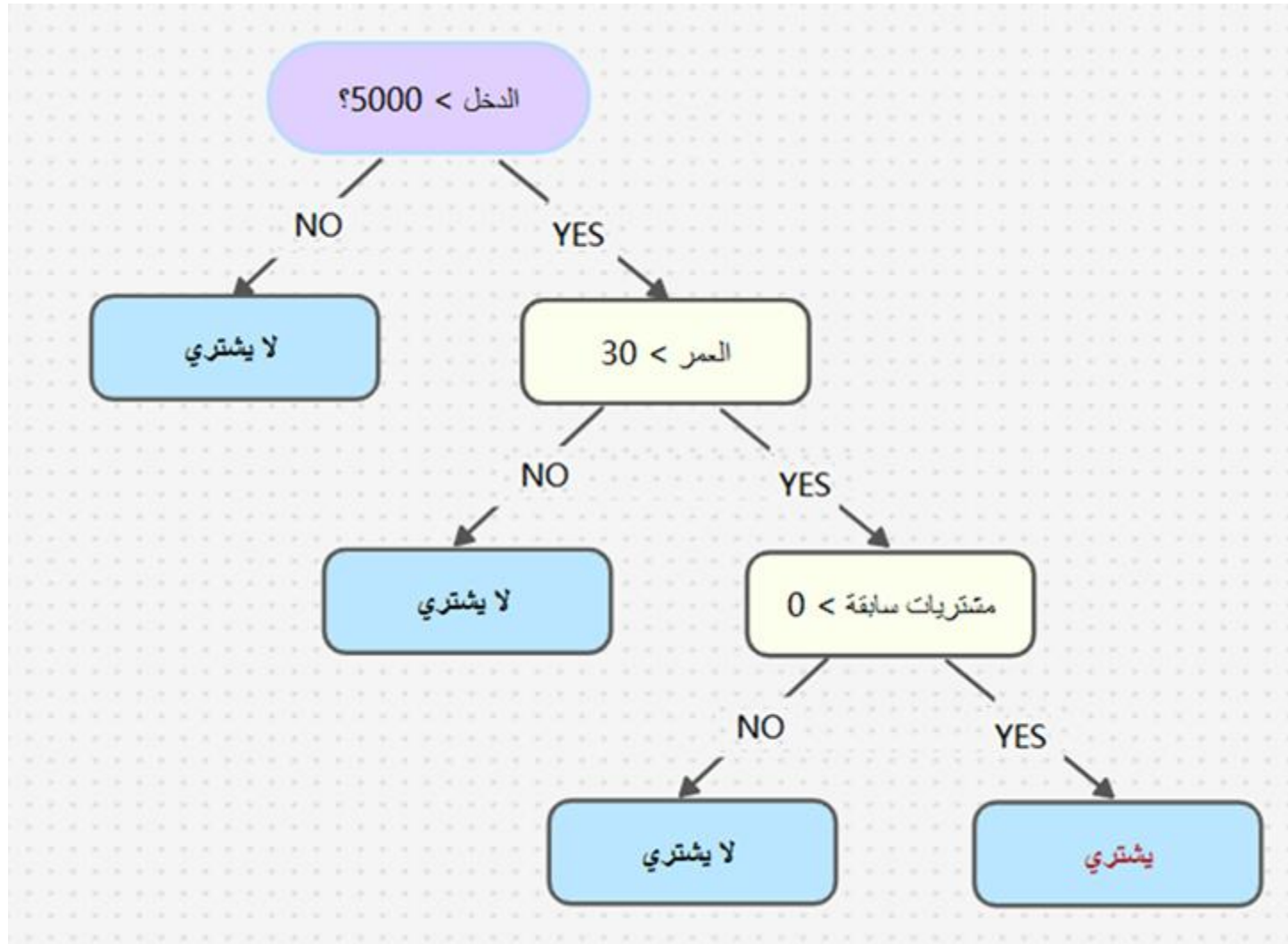
- إذا $\text{HasPrevPurchase} = \text{نعم}$ <-- C4, C6 كلهم يشتري.
- إذا $\text{HasPrevPurchase} = \text{لا}$ <-- C3, C8 كلهم لا يشتري.

إذن السؤال في العقدة الثالثة:

هل قام العميل بمشتريات سابقة ($0 <$)؟

هكذا أصبحت كل الأوراق (Leaves) نقية تمامًا (إما يشتري أو لا يشتري فقط)، وهذا ما تسعى إليه خوارزميات أشجار القرار عند التقسيم الجيد.

شجرة القرار للتصنيف (Classification Tree)



شجرة القرار للتصنيف (Classification Tree)

استخدام شجرة القرار للتنبؤ بحالة جديدة

مثال: عميل جديد:

- الدخل = 75000
- العمر = 34
- المشتريات السابقة = 2

خطوات المرور في الشجرة:

- الدخل < 50000 ؟ → نعم → نذهب لفرع اليمين.
- العمر < 30 ؟ → نعم → ننتقل للسؤال الثالث.
- المشتريات السابقة < 0 ؟ → نعم → نصل للورقة: **يشتري**.

إذن شجرة القرار تتنبأ بأن العميل الجديد سيشتري المنتج.

شجرة القرار (DECISION TREE)

شجرة قرار للتصنيف (Classification Tree)

- المخرج (Output) يكون فئة:
 - ناجح / راسب
 - مريض / سليم
 - نوع المنتج: A / B / C
- هذه الأشجار تستخدم لمهام التصنيف، حيث النتيجة منفصلة (Discrete)

شجرة قرار للانحدار (Regression Tree)

- المخرج يكون رقم حقيقي مستمر:
 - درجة الطالب من 100
 - سعر البيت
 - عدد الأيام المتوقعة لبقاء المريض في المستشفى
- هنا التنبؤ بقيمة رقمية مستمرة (Continuous)

شجرة القرار (DECISION TREE)

أثناء بناء الشجرة:

- في التصنيف: الخوارزمية تحاول أن تجعل كل فرع يحتوي على فئة واحدة قدر الإمكان، تستخدم مقاييس مثل Gini أو Entropy
- في الانحدار: تحاول أن تجعل الأفرع تحتوي على أرقام متقاربة، فتستخدم مقاييس مثل متوسط مربع الخطأ MSE لتقليل تشتت القيم داخل كل عقدة.

شجرة القرار للانحدار (Regression Tree)

مثال بسيط على شجرة قرار للانحدار (Regression)

نريد استخدام شجرة قرار للتنبؤ بدرجة الطالب في الامتحان (من 100) اعتمادًا على عدد ساعات الدراسة قبل الامتحان.
لدينا بيانات بسيطة لـ 6 طلاب:

الطالب	عدد ساعات الدراسة	الدرجة في الامتحان
A	2	50
B	3	60
C	4	65
D	6	80
E	7	85
F	8	90

شجرة القرار للانحدار (Regression Tree)

كيف تبني الشجرة

نبحث عن سؤال يقسم الطلاب لمجموعتين تكون درجات كل مجموعة متقاربة.
سؤال منطقي هنا:

هل عدد ساعات الدراسة ≤ 5 ساعات؟

- لو الجواب "لا" (أقل من 5 ساعات)

الطلاب: A و B و C

- الساعات: 2، 3، 4

- الدرجات: 50، 60، 65

- متوسط الدرجة ≈ 58

الشجرة تنتبأ لأي طالب جديد يدرس أقل من 5 ساعات بدرجة تقريبية 58.

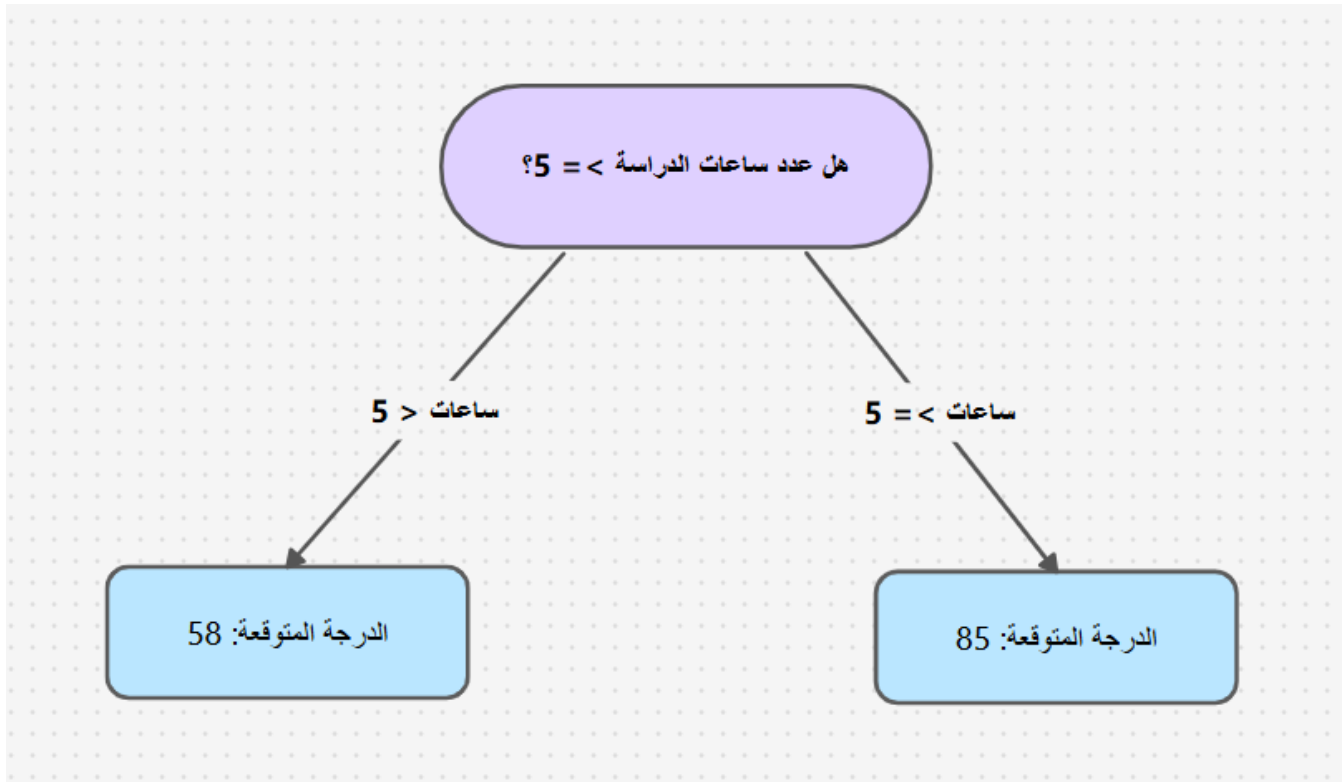
شجرة القرار للانحدار (Regression Tree)

- لو الجواب "نعم" (5 ساعات فأكثر)

الطلاب: D و E و F

- الساعات: 6، 7، 8
- الدرجات: 80، 85، 90
- متوسط الدرجة ≈ 85

الشجرة تتنبأ لأي طالب جديد يدرس 5 ساعات أو أكثر
بدرجة تقريبية **85**.



شجرة القرار للانحدار (Regression Tree)

كيف نستخدمها على طالب جديد؟

- طالب درس 3.5 ساعة:
• $3.5 > 5$ --> نذهب لفرع "لا" --> نتنبأ له تقريباً بـ 58.
- طالب درس 6.5 ساعة:
• $5 \leq 6.5$ --> نذهب لفرع "نعم" --> نتنبأ له تقريباً بـ 85.

إذا: شجرة الانحدار هنا لا تقول "ناجح/راسب"، بل تعطي رقم متوقع لدرجة الامتحان.