HIAST

| مقرر تعلم الآلة | السنة الرابعة – هندسة البرمجيات والذكاء الصنعي | المحاضرة الثانية |

أشجار القرار
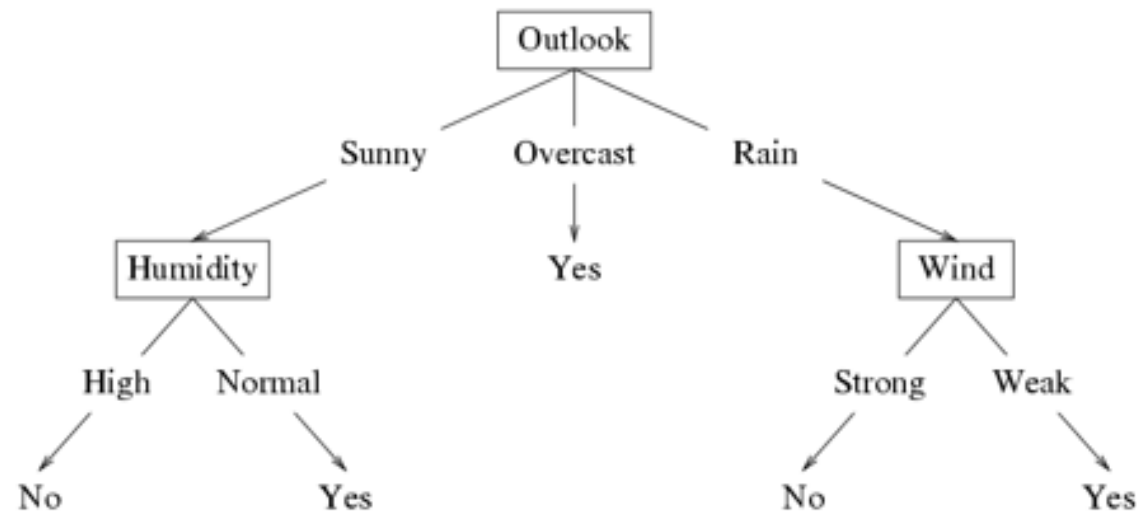# Decision Tree

د. رياض سنبل

Traditional ML Pipeline
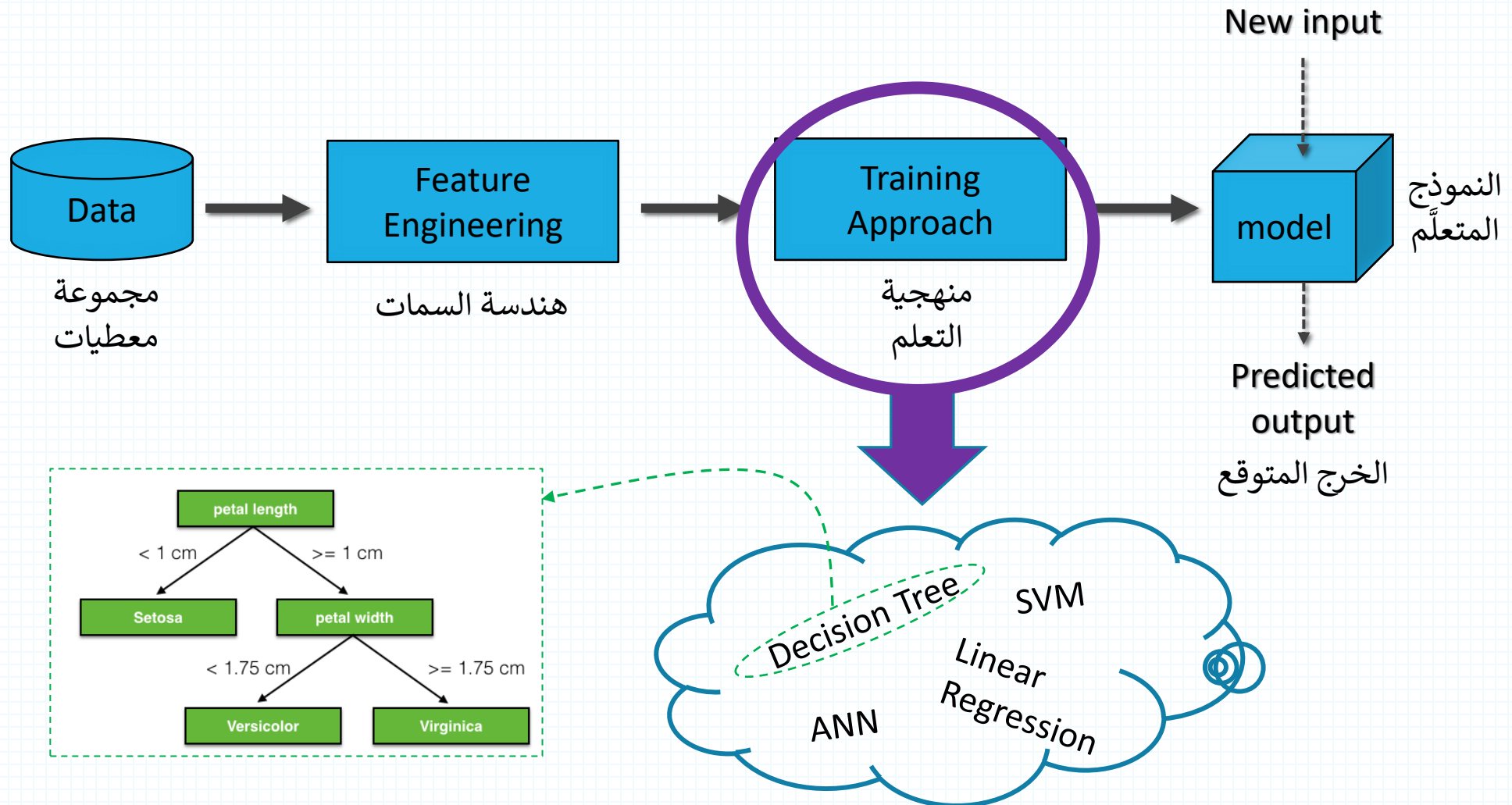
Example: Predict When This Player Will Play Tennis?

PlayTennis: training examples

| Day	Outlook	Temperature	Humidity	Wind	PlayTennis
D1	Sunny	Hot	High	Weak	No
D2	Sunny	Hot	High	Strong	No
D3	Overcast	Hot	High	Weak	Yes
D4	Rain	Mild	High	Weak	Yes
D5	Rain	Cool	Normal	Weak	Yes
D6	Rain	Cool	Normal	Strong	No
D7	Overcast	Cool	Normal	Strong	Yes
D8	Sunny	Mild	High	Weak	No
D9	Sunny	Cool	Normal	Weak	Yes
D10	Rain	Mild	Normal	Weak	Yes
D11	Sunny	Mild	Normal	Strong	Yes
D12	Overcast	Mild	High	Strong	Yes
D13	Overcast	Hot	Normal	Weak	Yes
D14	Rain	Mild	High	Strong	No

Decision Tree Hypothesis Space

- Internal nodes test the value of particular features xi and branch according to the results of the test.

- Leaf nodes specify the class.

Suppose the features are **Outlook** (x_1), **Temperature** (x_2), **Humidity** (x_3), and **Wind** (x_4). Then the feature vector $\mathbf{x} = (Sunny, Hot, High, Strong)$ will be classified as **No**. The **Temperature** feature is irrelevant.

If the features are Continuous…
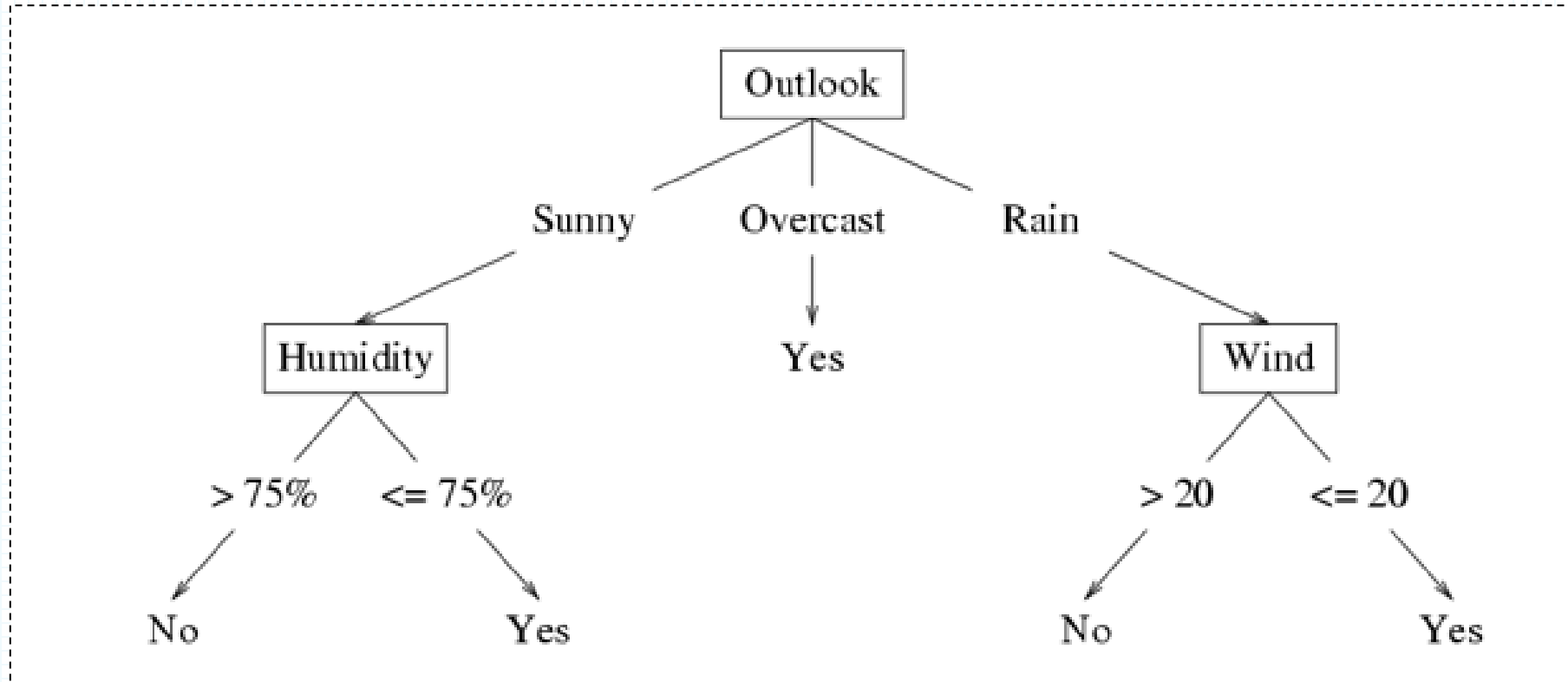
Decision Tree Decision Boundaries

- Decision trees divide the feature space into axis-parallel rectangles, and label each rectangle with one of the K class.
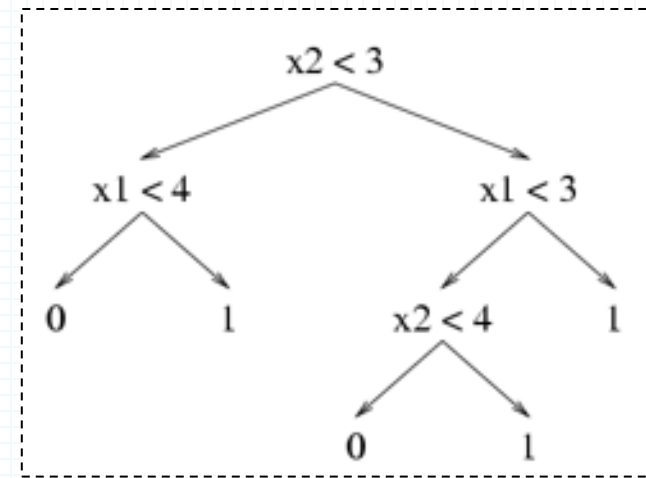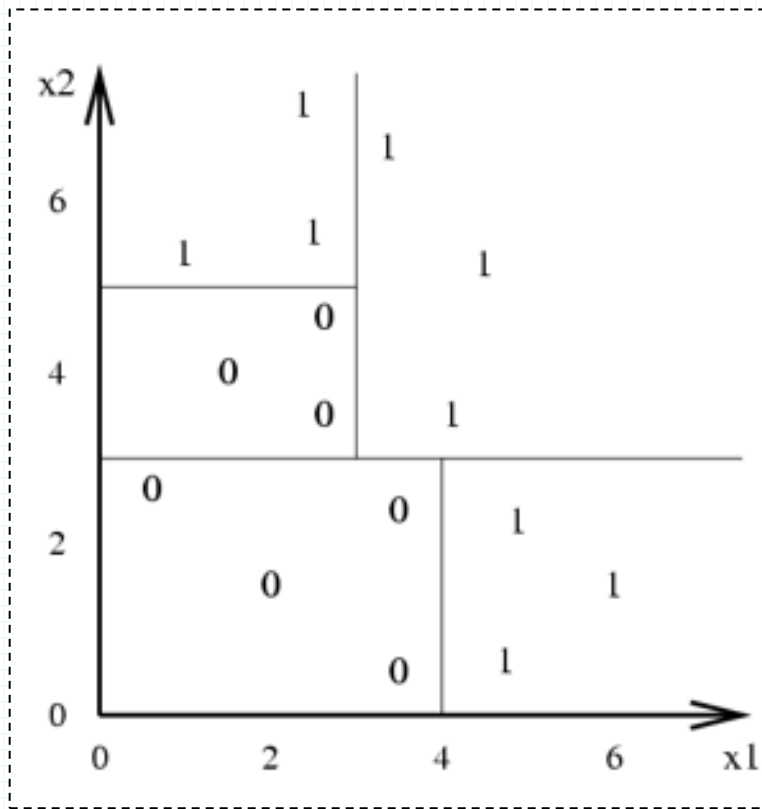
- Advantages:
  ◦ explicit relationship among features
  ◦ human interpretable model

- Limitations?

Learning Algorithm for Decision Trees

$$S = \{(\mathbf{x}_1, y_1), ..., (\mathbf{x}_N, y_N)\}$$

$$\mathbf{x} = (x_1, ..., x_d)$$

$$x_j, y \in \{0, 1\}$$

GrowTree(S)

if ($y = 0$ for all $\langle \mathbf{x}, y \rangle \in S$) **return** new leaf(0)

else if ($y = 1$ for all $\langle \mathbf{x}, y \rangle \in S$) **return** new leaf(1)

else

choose best attribute x_j

$S_0 = $ all $\langle \mathbf{x}, y \rangle \in S$ with $x_j = 0$;

$S_1 = $ all $\langle \mathbf{x}, y \rangle \in S$ with $x_j = 1$;

return new node(x_j, GrowTree(S_0), GrowTree(S_1))