

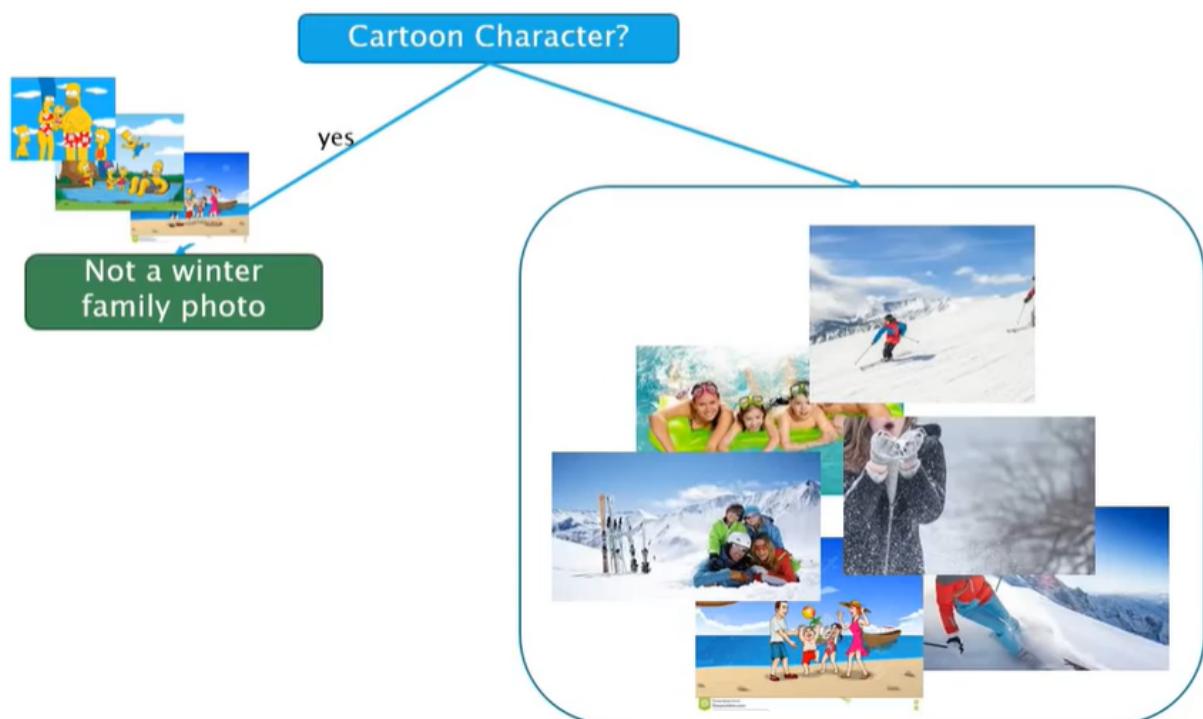
Decision Tree

1. How Decision Tree works

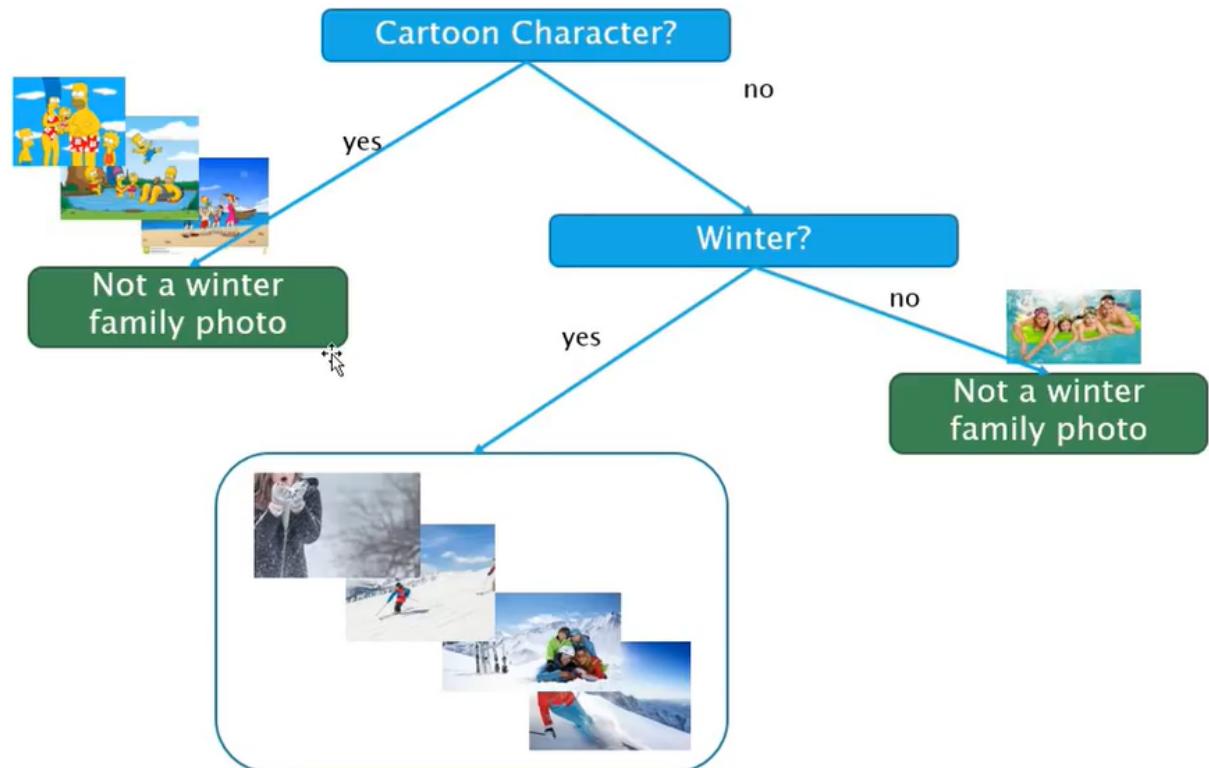
- How do you teach your baby to pick winter family vacation photo?
- You may say to consider snow, human, and the number of members



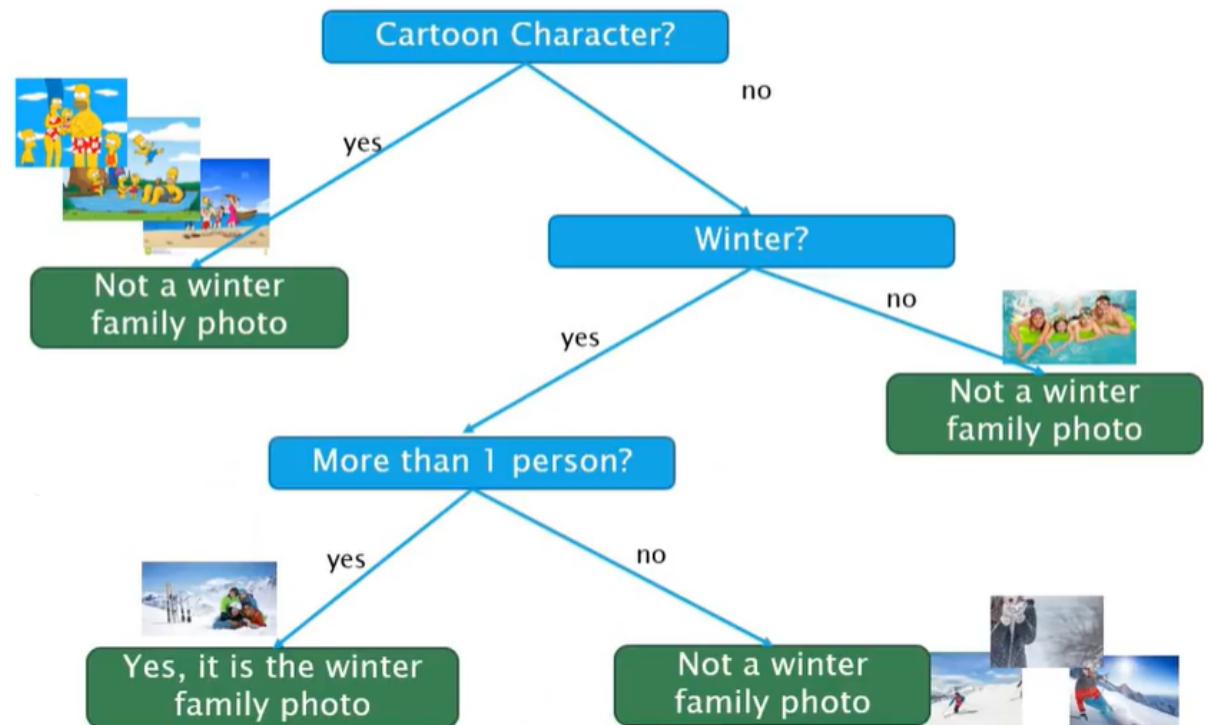
- We definitely know the cartoon is not a family photo



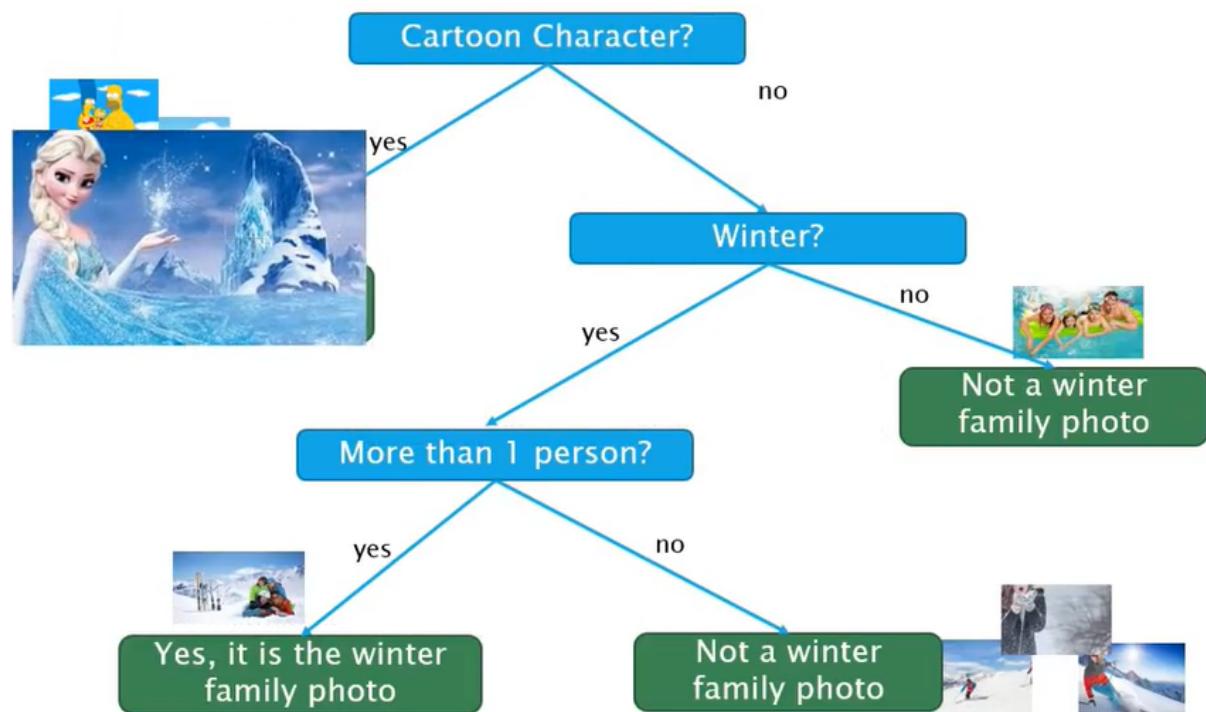
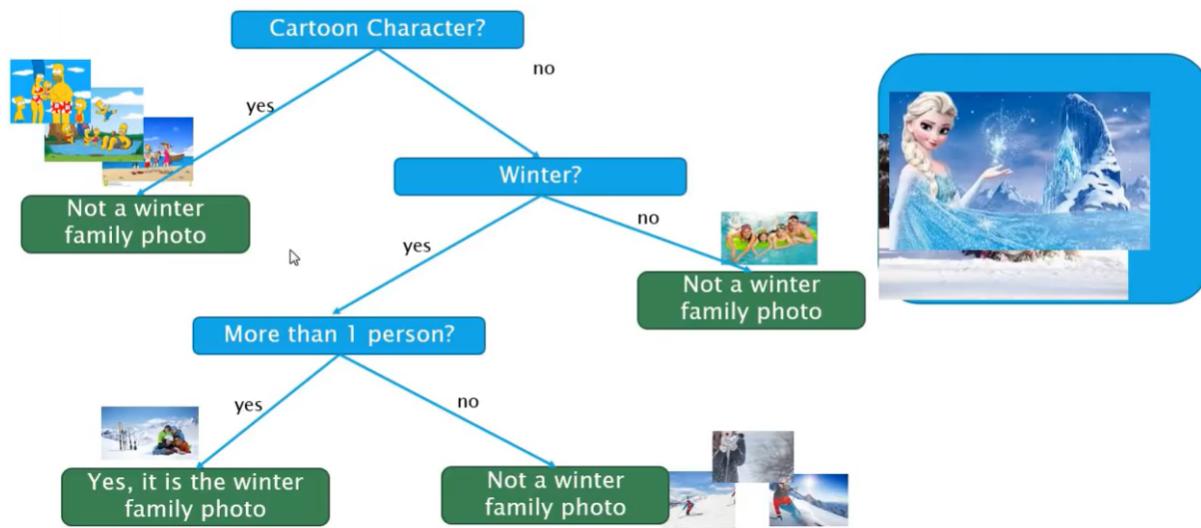
- We definitely know summer is not a winter photo

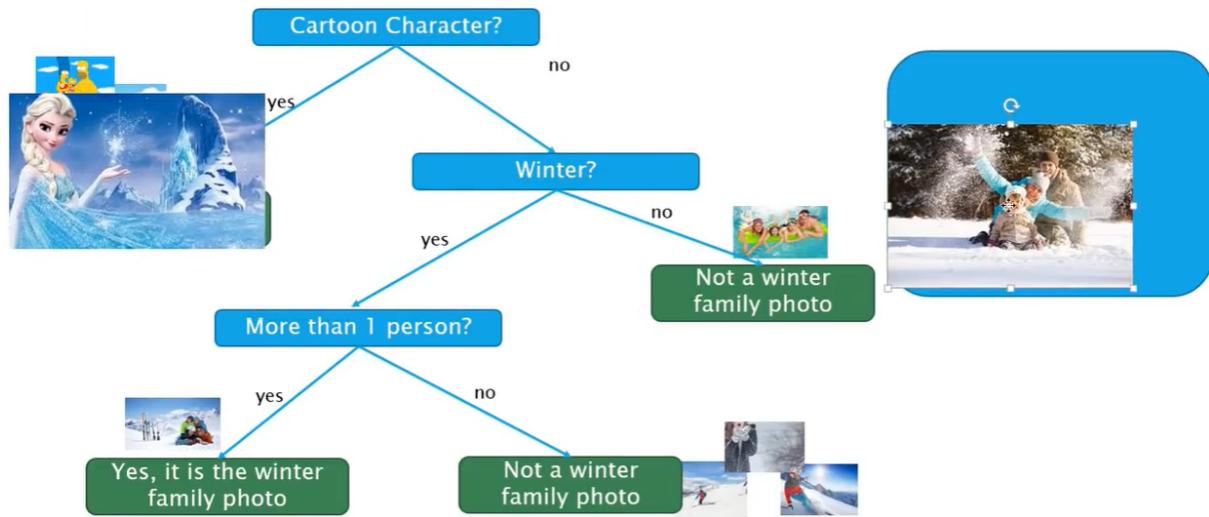


- We know family photo has more than 1 person

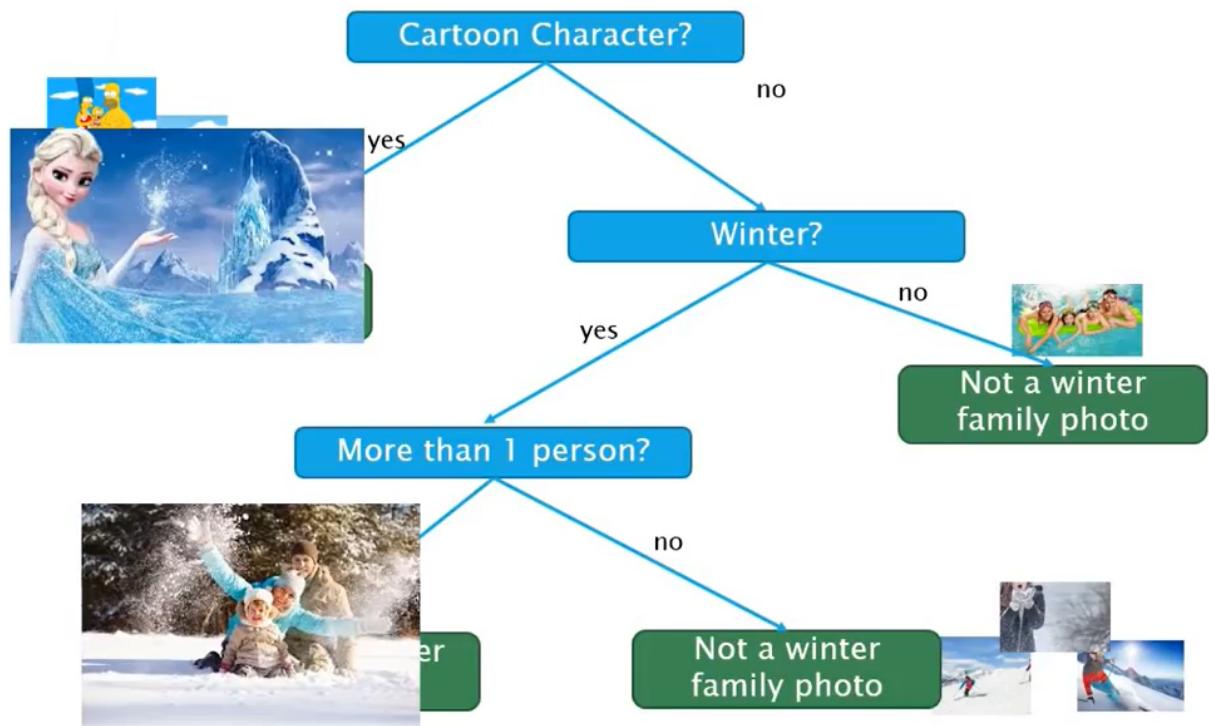


- Now, your baby can recognize more pictures using this decision tree





- Yes, the baby is the machine, and the machine learned from you



2. Decision Tree, two main concepts

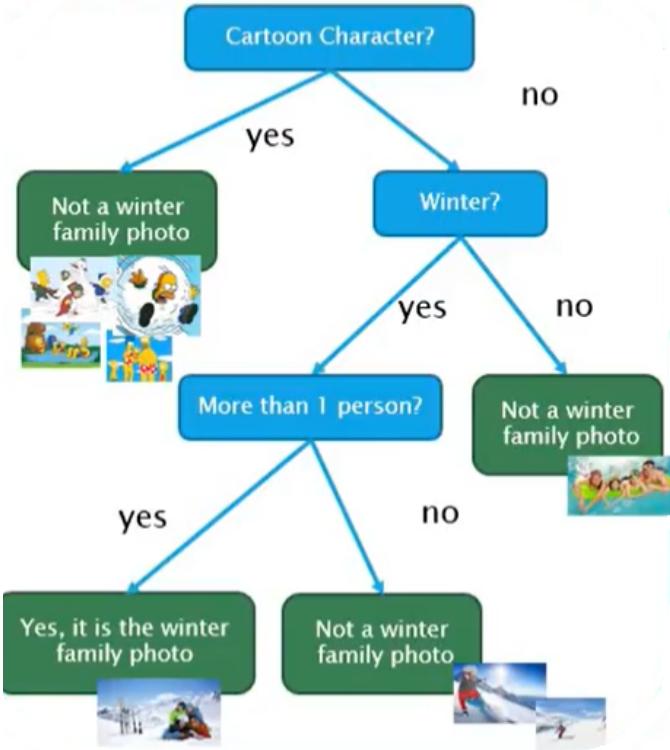
- Define Problem

- Collect training data
- Extract Data
- **Build a tree**
- Deploy machine
- **Test with test data**

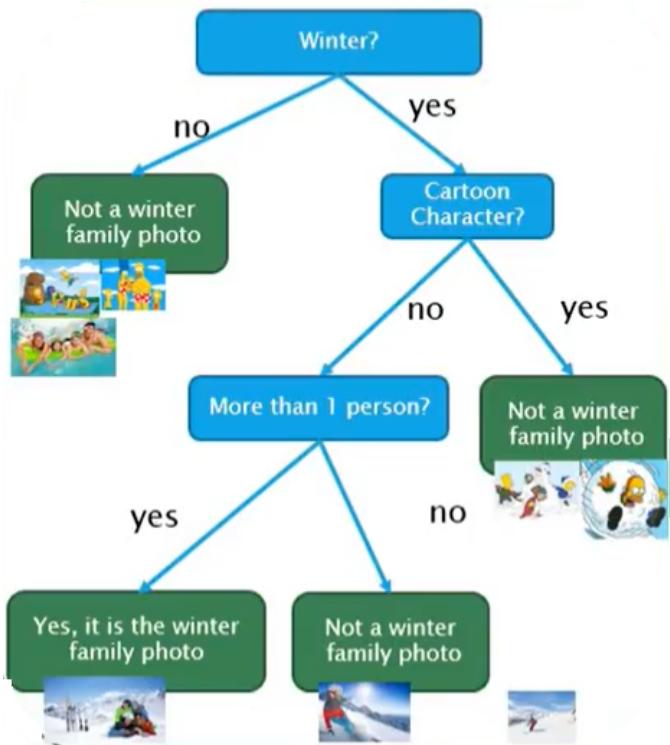
(1) Build a Decision Tree - Extract training data

img	cartoon	winter	> 1	Family winter photo
	No	Yes	Yes	Yes
	No	Yes	No	No
	Yes	No	Yes	No
	Yes	Yes	Yes	No
	No	Yes	No	No
	No	No	Yes	No
	Yes	No	Yes	No
	yes	yes	no	no

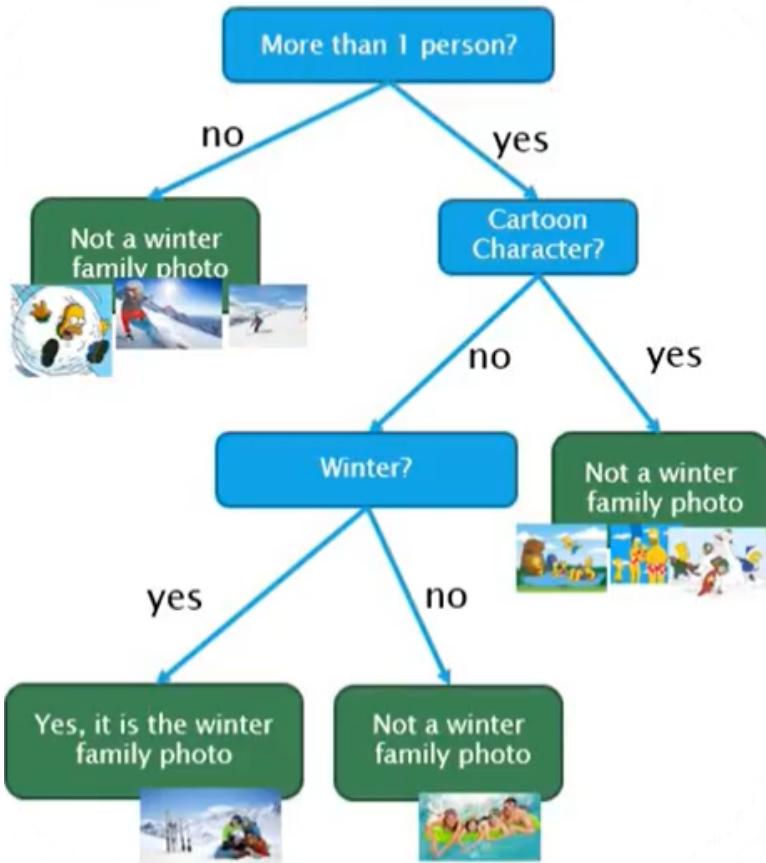
- We had 8 pictures, Now just 4 left



- We had 8 pictures, Now just 5 left

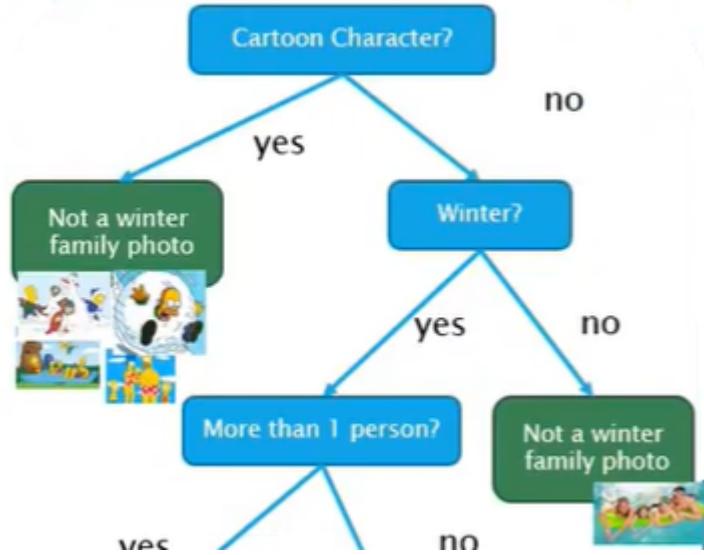


- We had 8 pictures, Now just 5 left

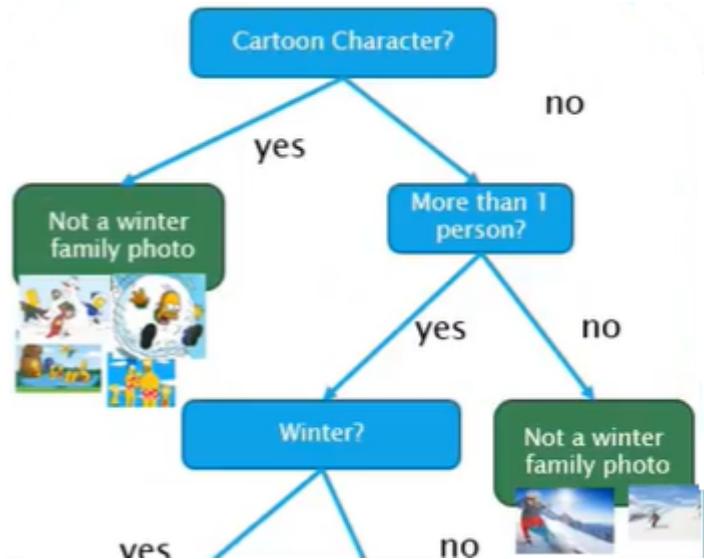


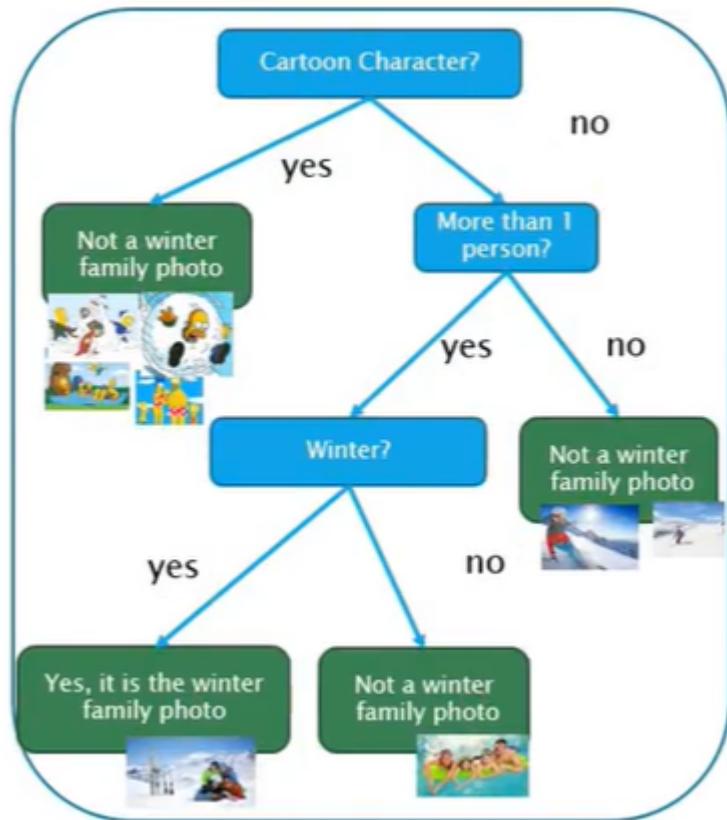
- **Choose best feature to split**

- We had 5 pictures, Now 4 left



- We had 5 pictures, Now 3 left

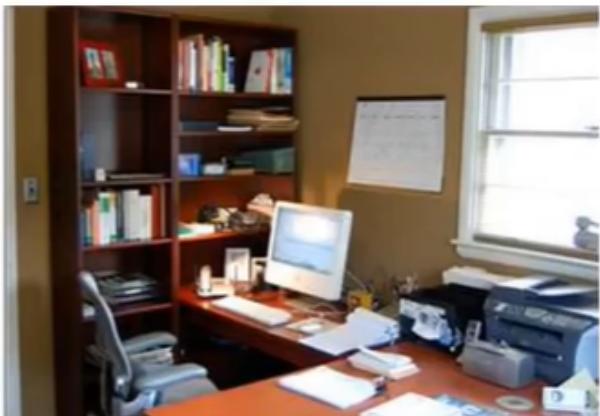




※ entropy



High Entropy (messy)

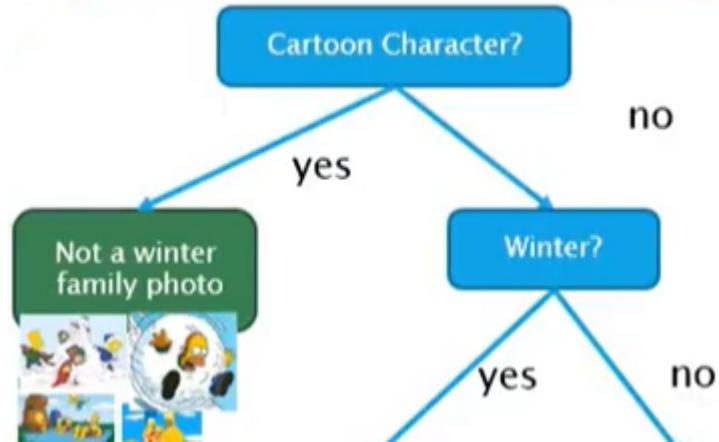


Low Entropy (Clean)

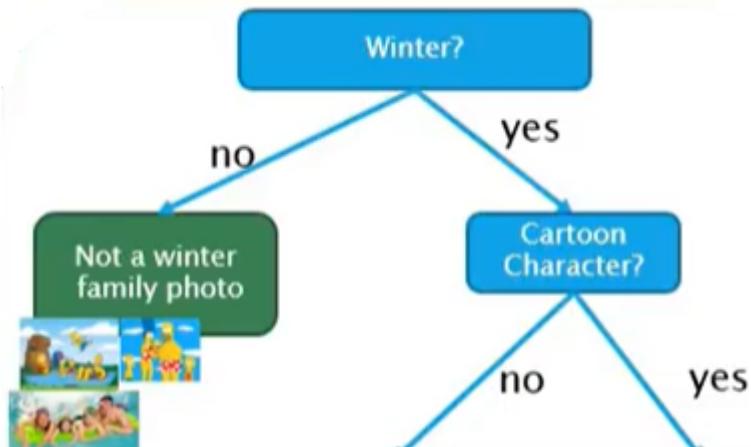
※ Information Gain (base entropy - new entropy)

- We had 8 pictures, Now 4 left

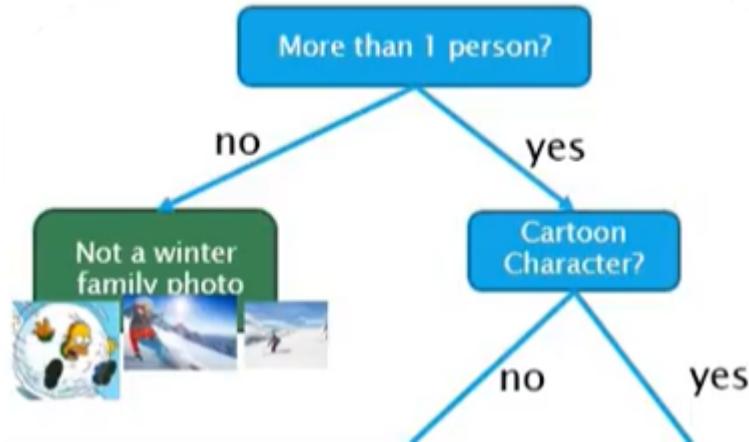
- Information Gain : 4
- Highest information gain



- We had 8 pictures, Now 5 left
 - Information Gain : 3



- We had 8 pictures, Now 5 left
 - Information Gain : 3



3. ID3 Algorithm (Entropy and Information Gain)

- Calculate Entropy to find family photo of 8 photos

img	cartoon	winter	> 1	Family winter photo
	No	Yes	Yes	Yes
	No	Yes	No	No
	Yes	No	Yes	No
	Yes	Yes	Yes	No
	No	Yes	No	No
	No	No	Yes	No
	Yes	No	Yes	No
	yes	yes	no	no

Total 8 photos

1 photo winter family photo

7 photo **not** winter family photo

$$= \text{Entropy}([1+, 7-])$$

$$= -(1/8) * \log(1/8) - (7/8) * \log(7/8) = 0.543$$

※ Entropy = $- p(+)^* \log(p(+)) - p(-)^* \log(p-)$

- **Information Gain(decrease in entropy after a dataset is split on an attribute)**

Information Gain(winter family photo, cartoon)

$$= E(\text{winter family photo}) - E(\text{winter family photo, cartoon})$$

$$= 0.543 - (4/8 * E([0+, 4-]) + 4/8 * E([1+, 3-]))$$

$$= 0.138$$

Information Gain(winter family photo, winter)

$$= E(\text{winter family photo}) - E(\text{winter family photo, winter})$$

$$= 0.543 - (5/8 * E([1+, 4-]) + 3/8 * E([0+, 3-]))$$

$$= 0.093$$

Information Gain(winter family photo, >1)

$$= E(\text{winter family photo}) - E(\text{winter family photo, } >1)$$

$$= 0.543 - (5/8 * E([1+, 4-]) + 3/8 * E([0+, 3-]))$$

$$= 0.093$$