

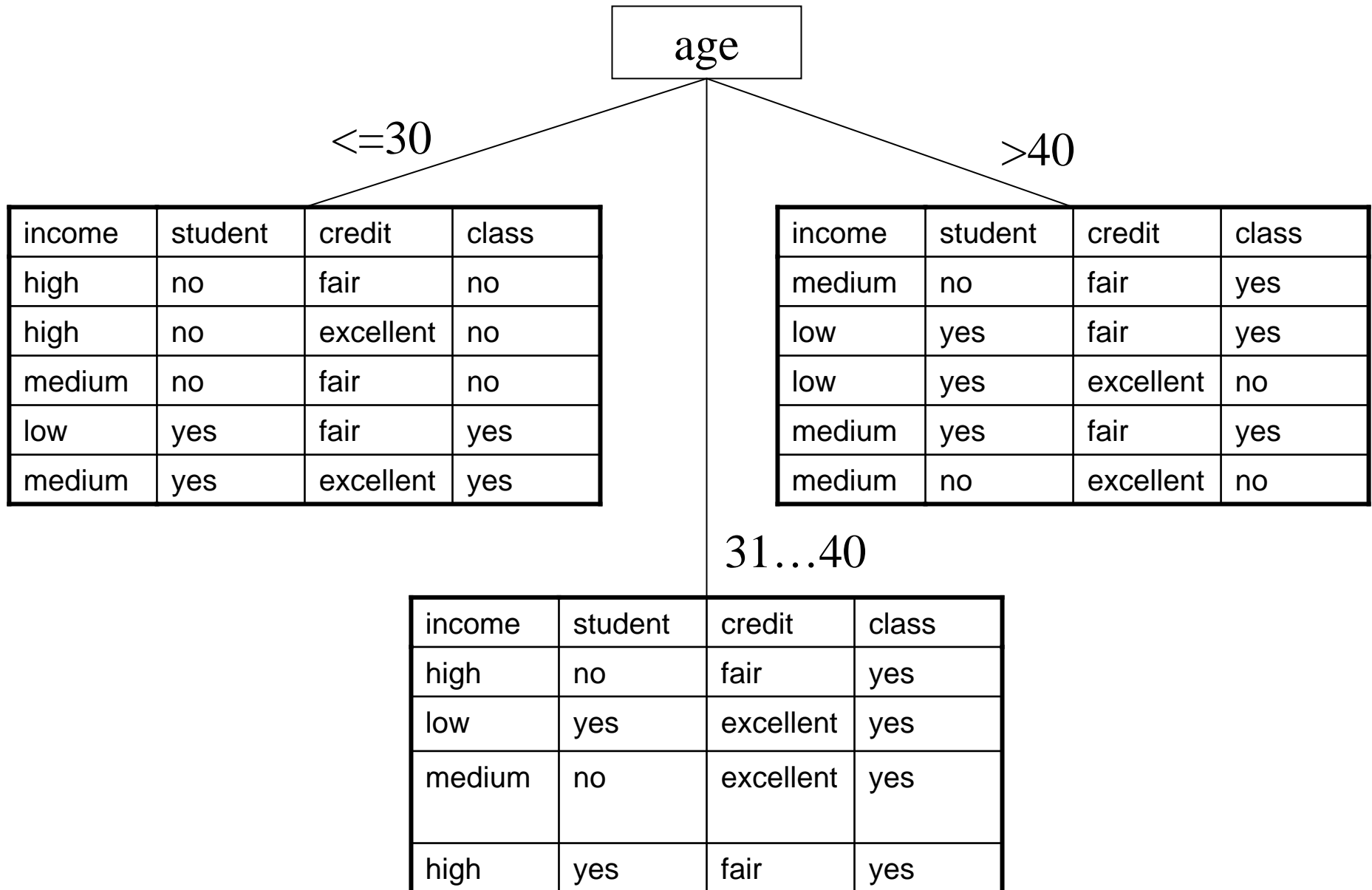
Building the Trees with Special Majority Voting

- **Examples**
- Professor Anita Wasilewska

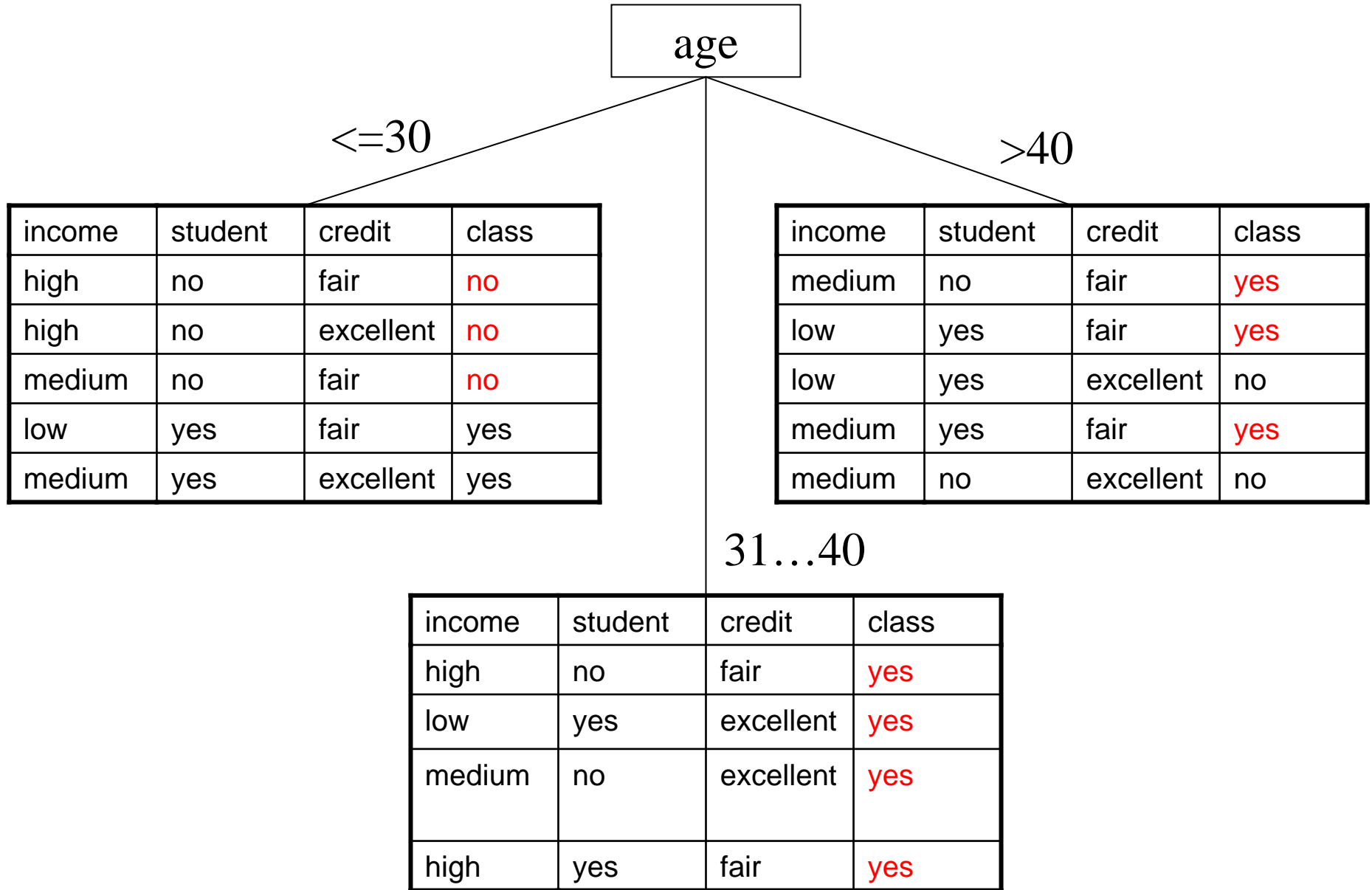
Training Dataset

rec	Age	Income	Student	Credit_rating	Buys_computer (CLASS)
r1	<=30	High	No	Fair	No
r2	<=30	High	No	Excellent	No
r3	31...40	High	No	Fair	Yes
r4	>40	Medium	No	Fair	Yes
r5	>40	Low	Yes	Fair	Yes
r6	>40	Low	Yes	Excellent	No
r7	31...40	Low	Yes	Excellent	Yes
r8	<=30	Medium	No	Fair	No
r9	<=30	Low	Yes	Fair	Yes
r10	>40	Medium	Yes	Fair	Yes
r11	<=30	Medium	Yes	Excellent	Yes
r12	31...40	Medium	No	Excellent	Yes
r13	31...40	High	Yes	Fair	Yes
r14	>40	Medium	No	Excellent	No

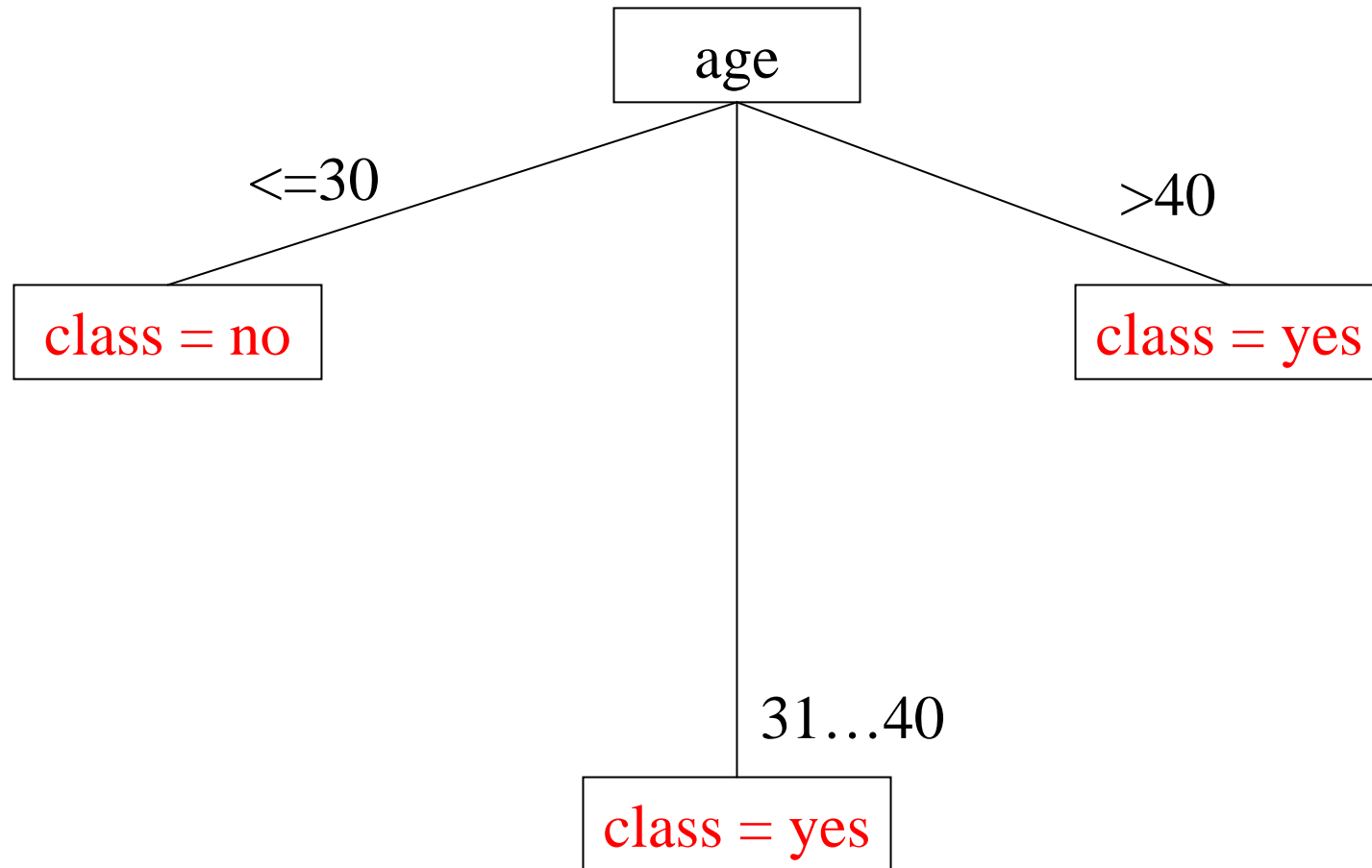
Building The Tree: "age" as root



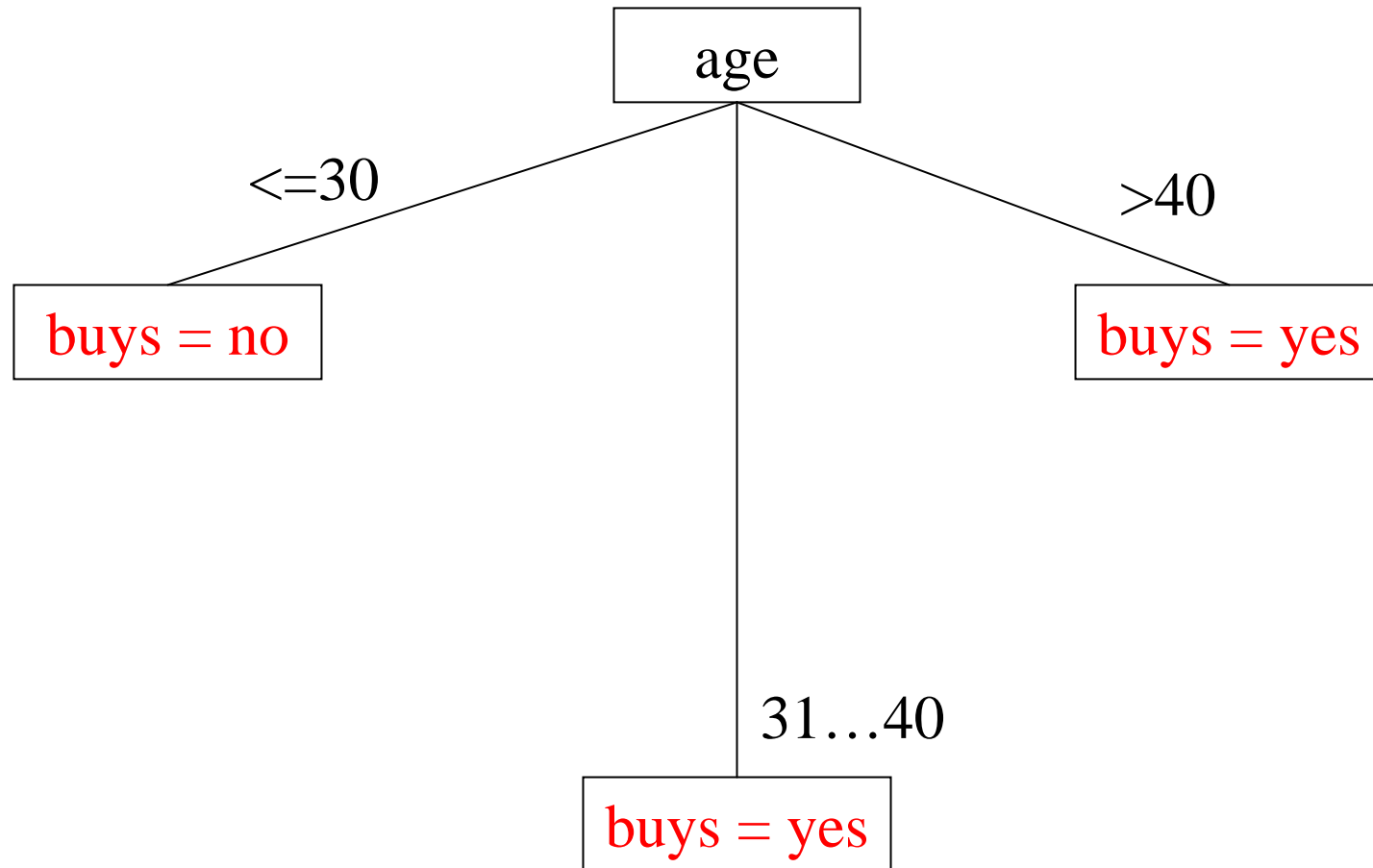
Majority Voting: all branches



Majority Voting: all branches



Finished Tree for class = "buys"



Building The Tree: “credit” as root

credit

fair

excellent

age	income	student	class
<=30	high	no	no
31...40	high	no	yes
>40	medium	no	yes
>40	low	yes	yes
<=30	medium	no	no
<=30	low	yes	yes
>40	medium	yes	yes
31...40	high	yes	yes

age	income	student	class
<=30	high	no	no
>40	low	yes	no
31...40	low	yes	yes
<=30	medium	yes	yes
31...40	medium	no	yes
>40	medium	no	no

Majority Voting: “fair” branch

credit

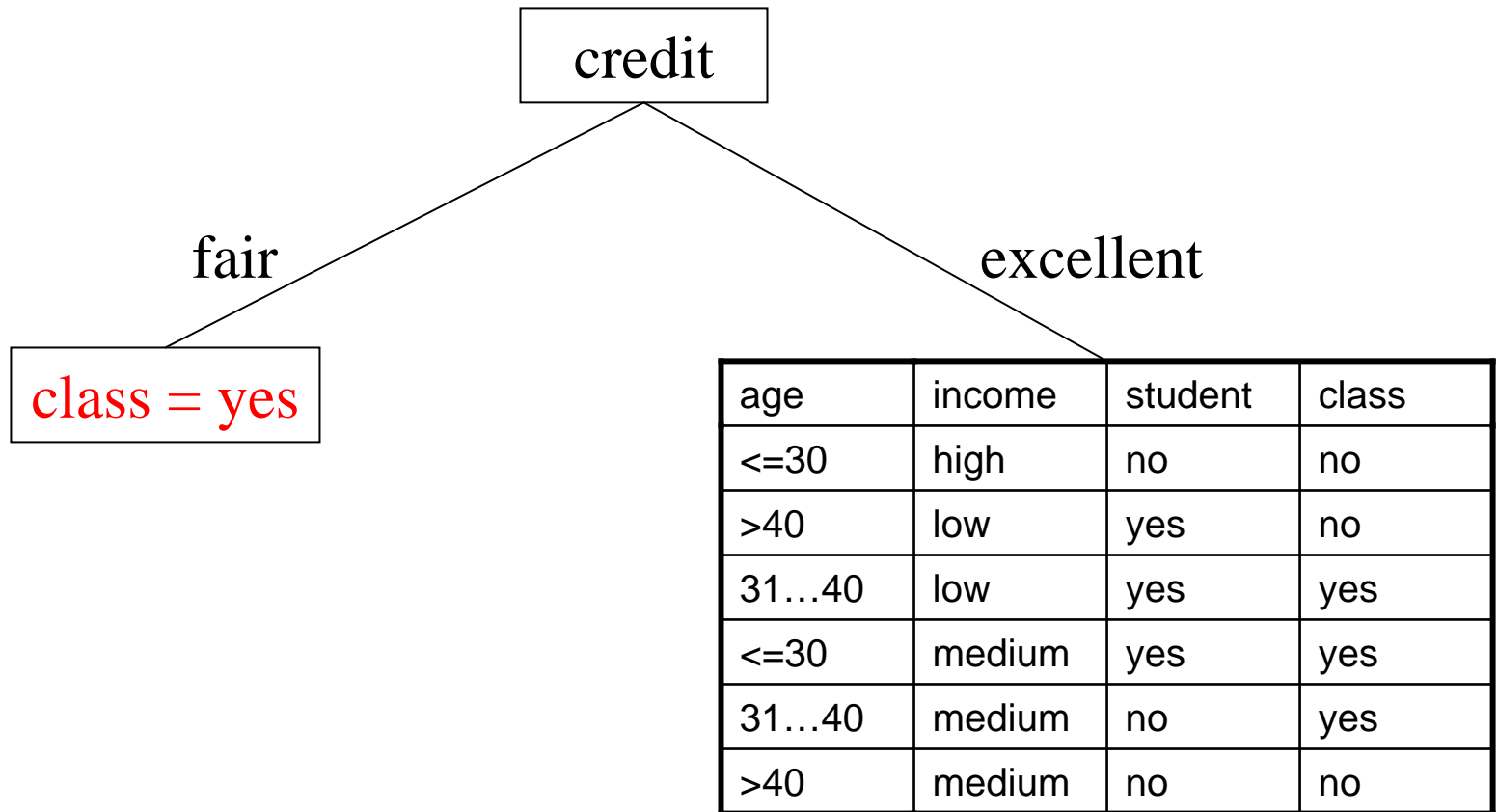
fair

excellent

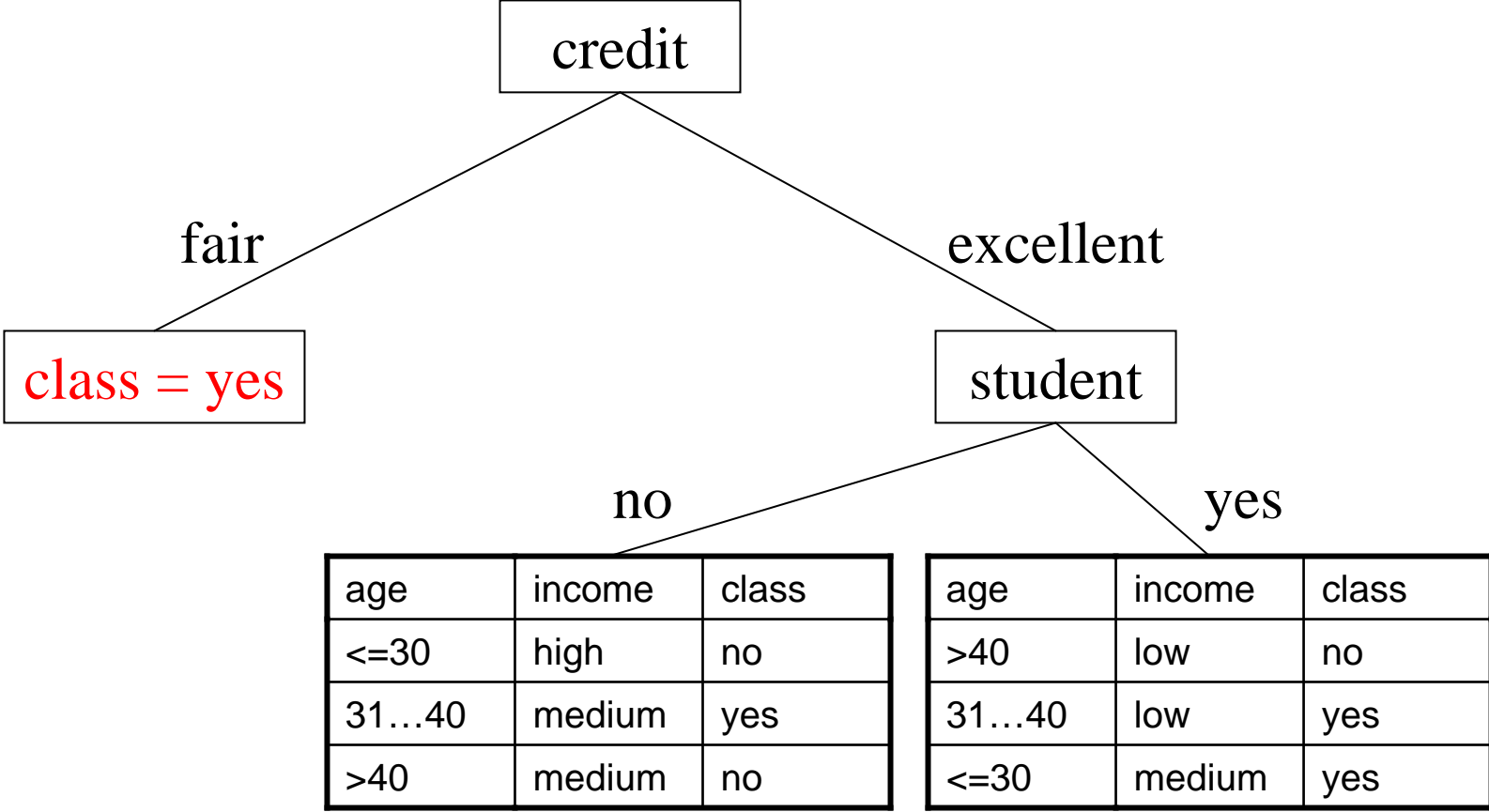
age	income	student	class
<=30	high	no	no
31...40	high	no	yes
>40	medium	no	yes
>40	low	yes	yes
<=30	medium	no	no
<=30	low	yes	yes
>40	medium	yes	yes
31...40	high	yes	yes

age	income	student	class
<=30	high	no	no
>40	low	yes	no
31...40	low	yes	yes
<=30	medium	yes	yes
31...40	medium	no	yes
>40	medium	no	no

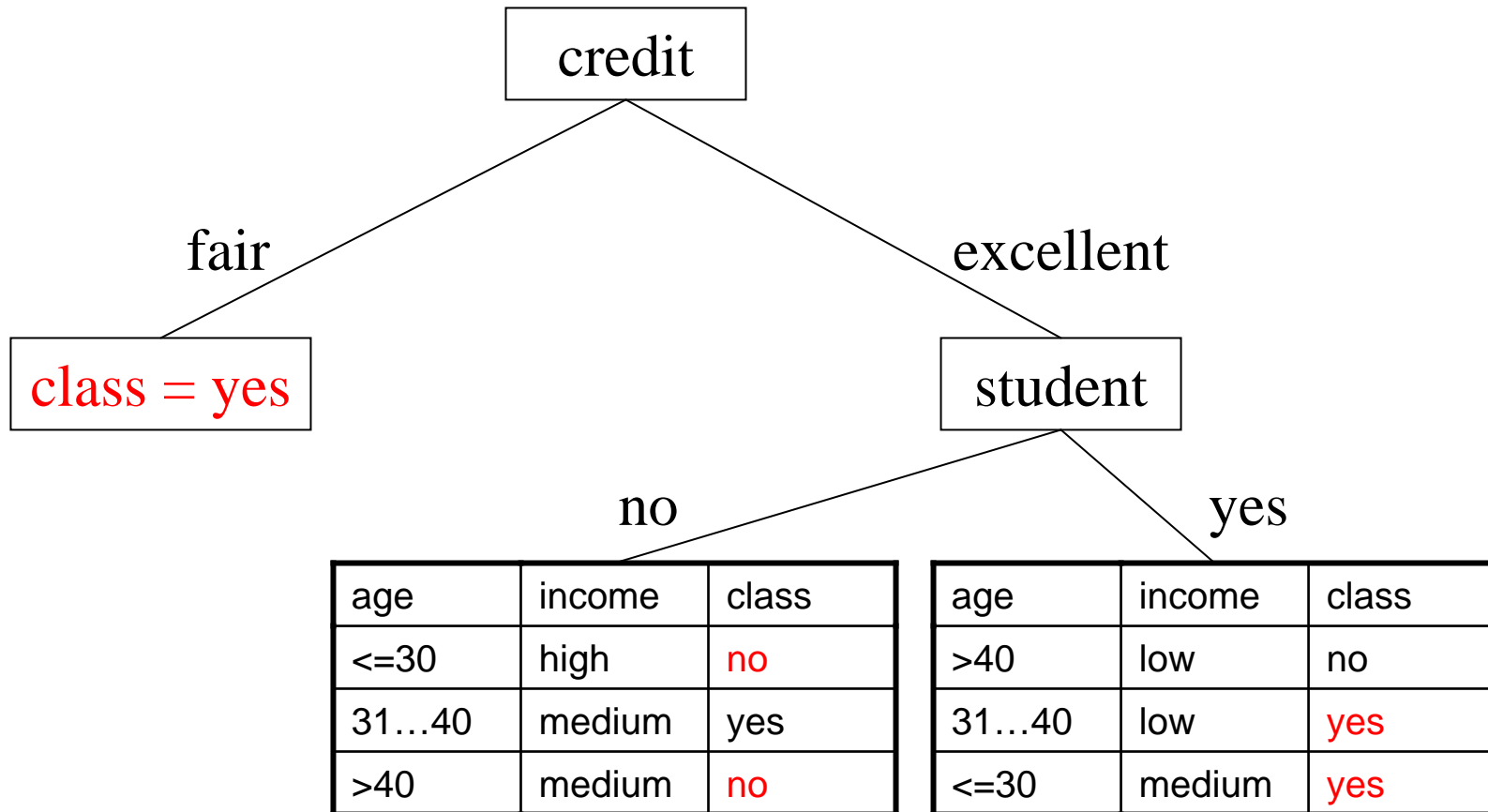
Majority Voting: “fair” branch



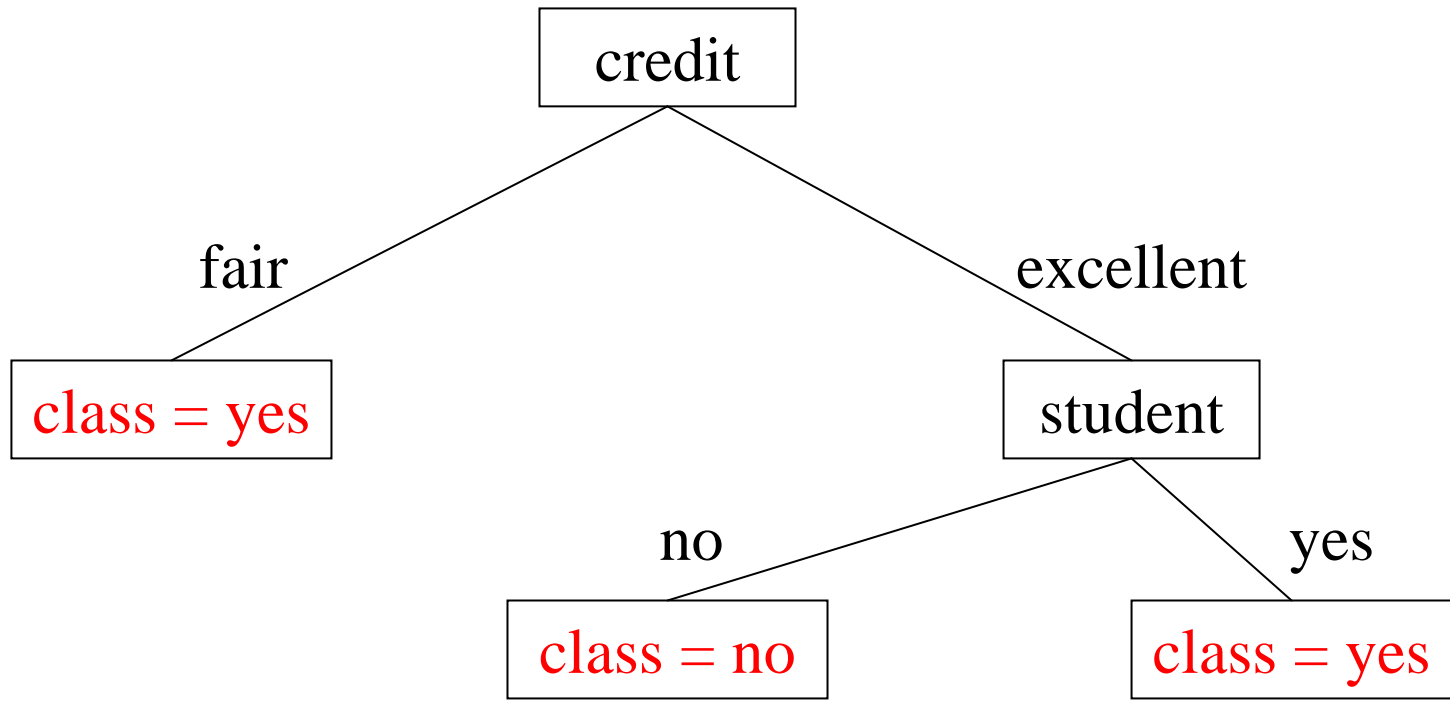
Building The Tree: we choose "student" on "excellent" branch



Majority Voting: “no” and “yes” branches

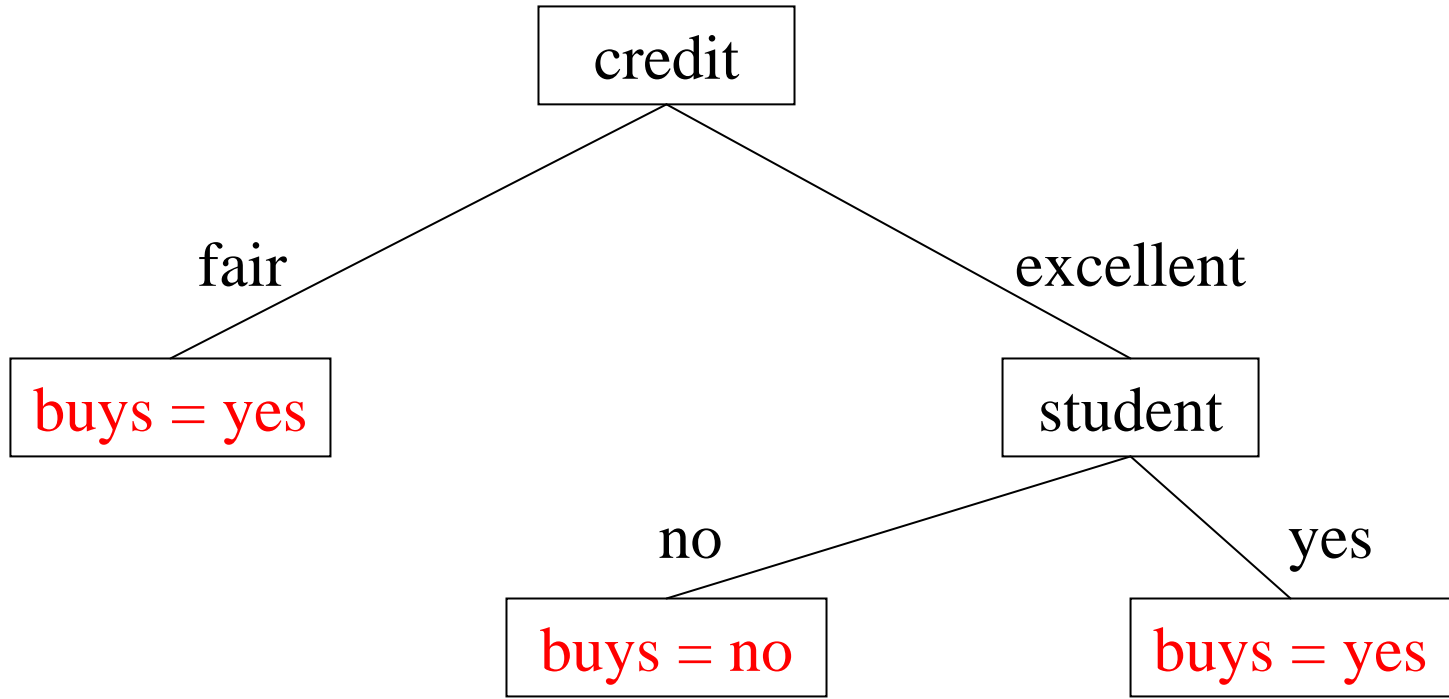


Majority Voting: “no” and “yes” branches



Finished Tree (with majority vote) for

class = "buys"



Exercise

- For all examples on previous slides,
- (1) Write rules determined by the tree in the predicate form
- (2) Compute rules accuracy (with respect of the training dataset)
- (3) Compute Predictive Accuracy with respect of the TEST dataset on the next slide.

Test Data Set

rec	Age	Income	Student	Credit_rating	Buys_computer(CLASS)
r1	<=30	Low	yes	Fair	no
r2	<=30	High	yes	Excellent	no
r3	<=30	High	yes	Fair	yes
r4	31...40	Medium	yes	Fair	no
r5	>40	Low	yes	Fair	yes
r6	>40	Low	no	Excellent	yes
r7	31...40	High	yes	Excellent	yes
r8	<=30	Medium	No	Fair	no
r9	31...40	Low	no	Excellent	yes
r10	>40	Medium	no	Fair	no