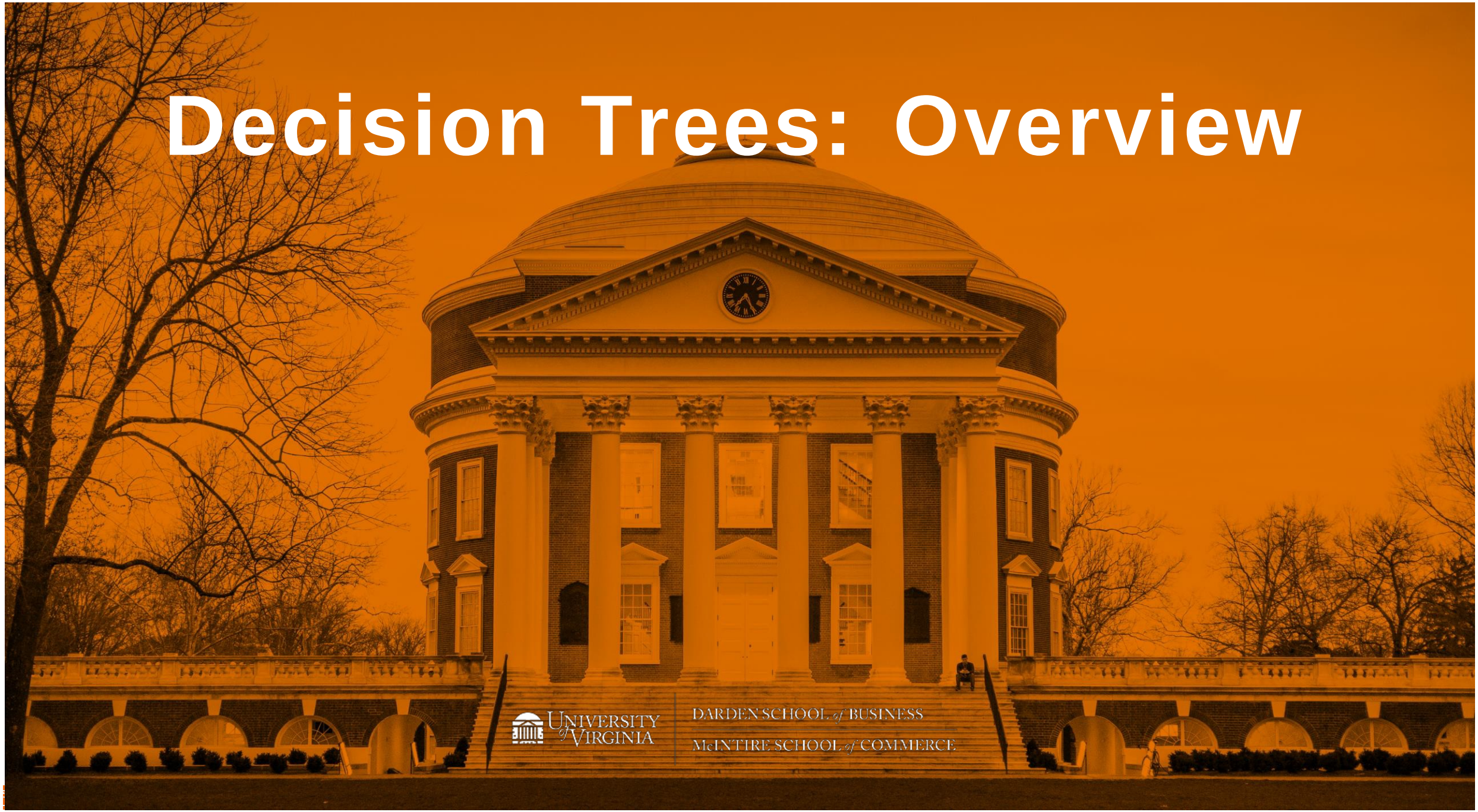
Decision Trees: Overview

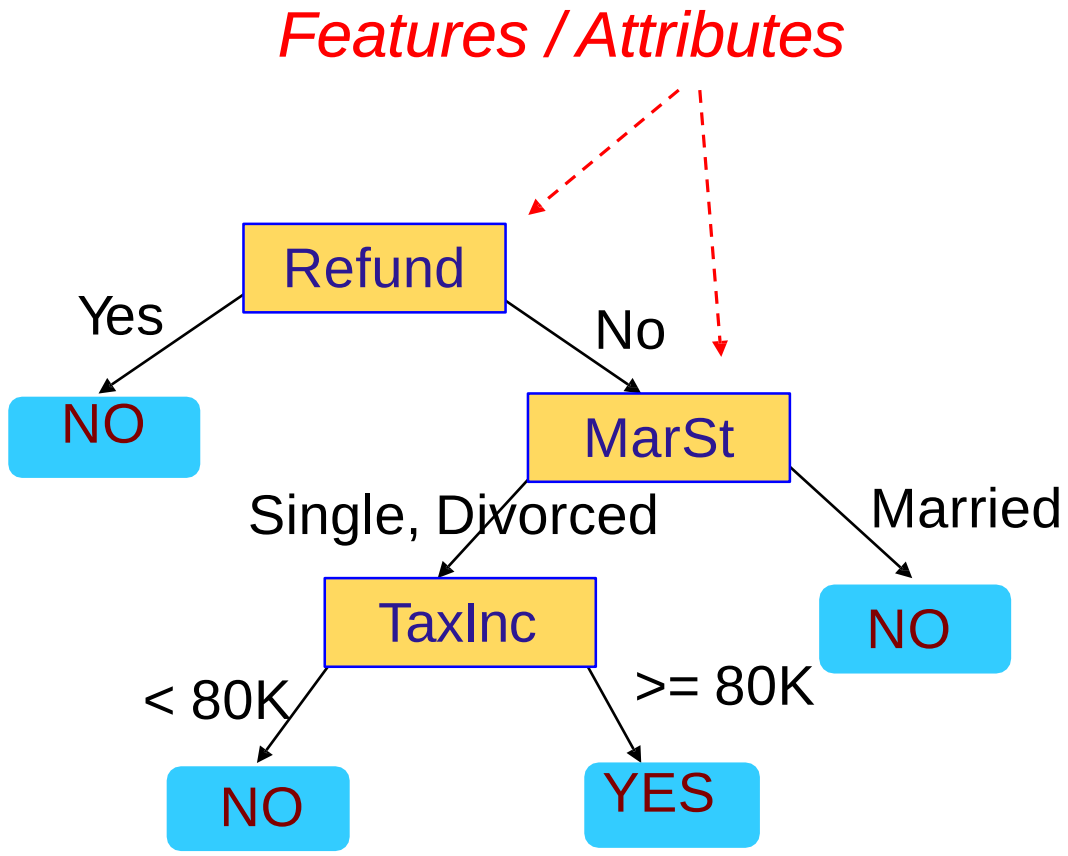


DECISION TREE EXAMPLE

DECISION TREE EXAMPLE

class

Tid	Refund	Marital Status	Taxable Income	Fraud
1	Yes	Single	125K	NO
2	No	Married	100K	NO
3	No	Single	70K	NO
4	Yes	Married	120K	NO
5	No	Divorced	95K	YES
6	No	Married	60K	NO
7	Yes	Divorced	220K	NO
8	No	Single	85K	YES
9	No	Married	75K	NO
10	No	Single	90K	YES



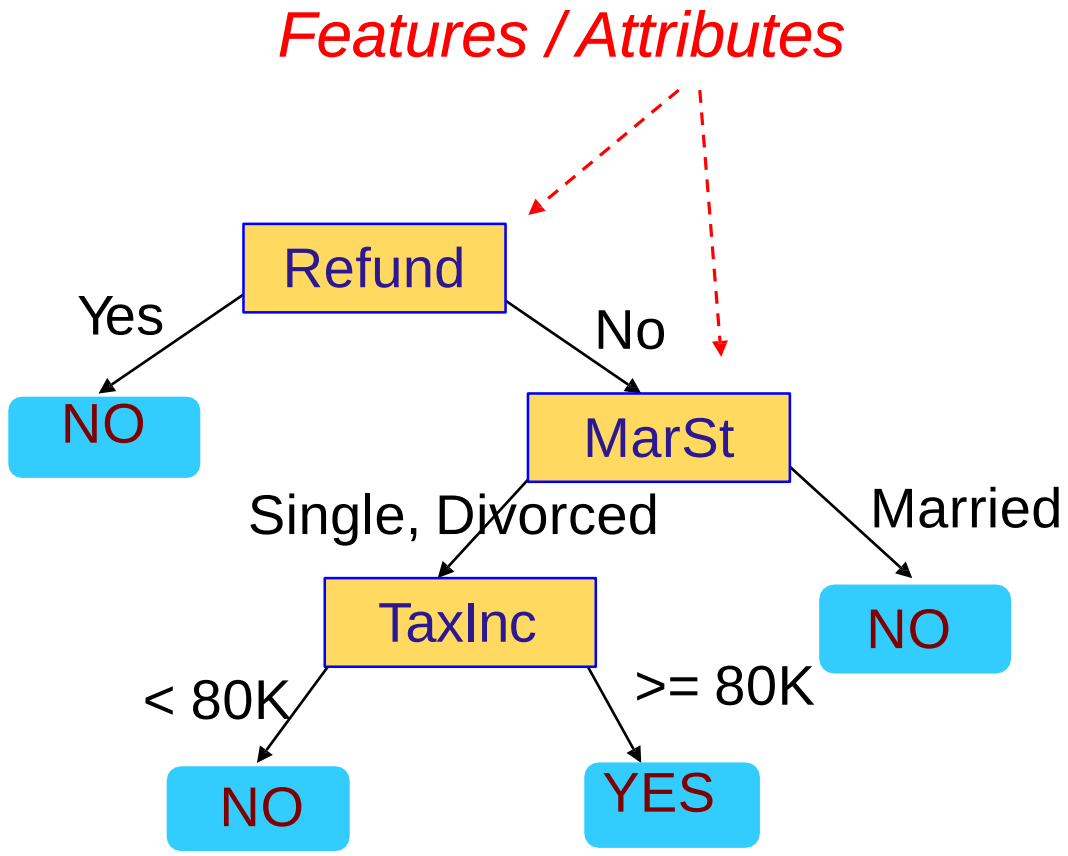
Training Data Matrix

Decision Tree Model

DECISION TREE ILLUSTRATION

class

Tid	Refund	Marital Status	Taxable Income	Fraud
1	Yes	Single	125K	NO
2	No	Married	100K	NO
3	No	Single	70K	NO
4	Yes	Married	120K	NO
5	No	Divorced	95K	YES
6	No	Married	60K	NO
7	Yes	Divorced	220K	NO
8	No	Single	85K	YES
9	No	Married	75K	NO
10	No	Single	90K	YES



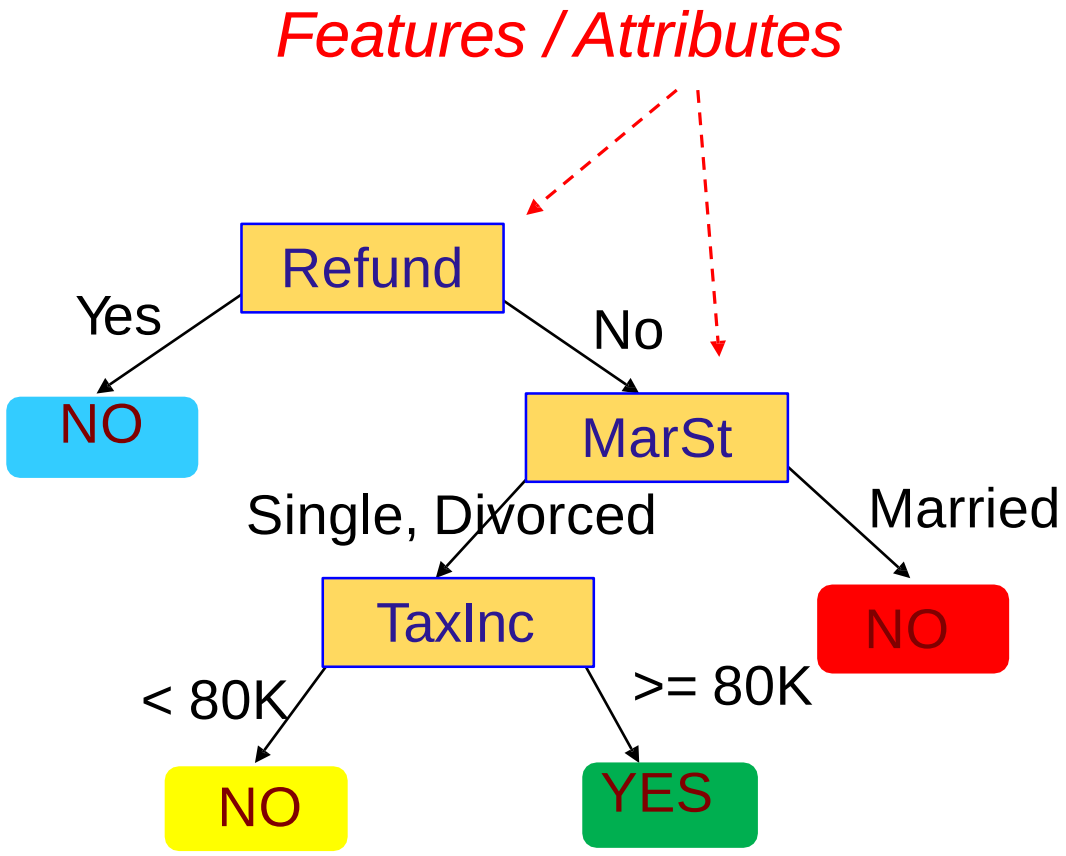
Training Data Matrix

Decision Tree Model

DECISION TREE TRAINING FIT

class

Tid	Refund	Marital Status	Taxable Income	Fraud
1	Yes	Single	125K	NO
2	No	Married	100K	NO
3	No	Single	70K	NO
4	Yes	Married	120K	NO
5	No	Divorced	95K	YES
6	No	Married	60K	NO
7	Yes	Divorced	220K	NO
8	No	Single	85K	YES
9	No	Married	75K	NO
10	No	Single	90K	YES



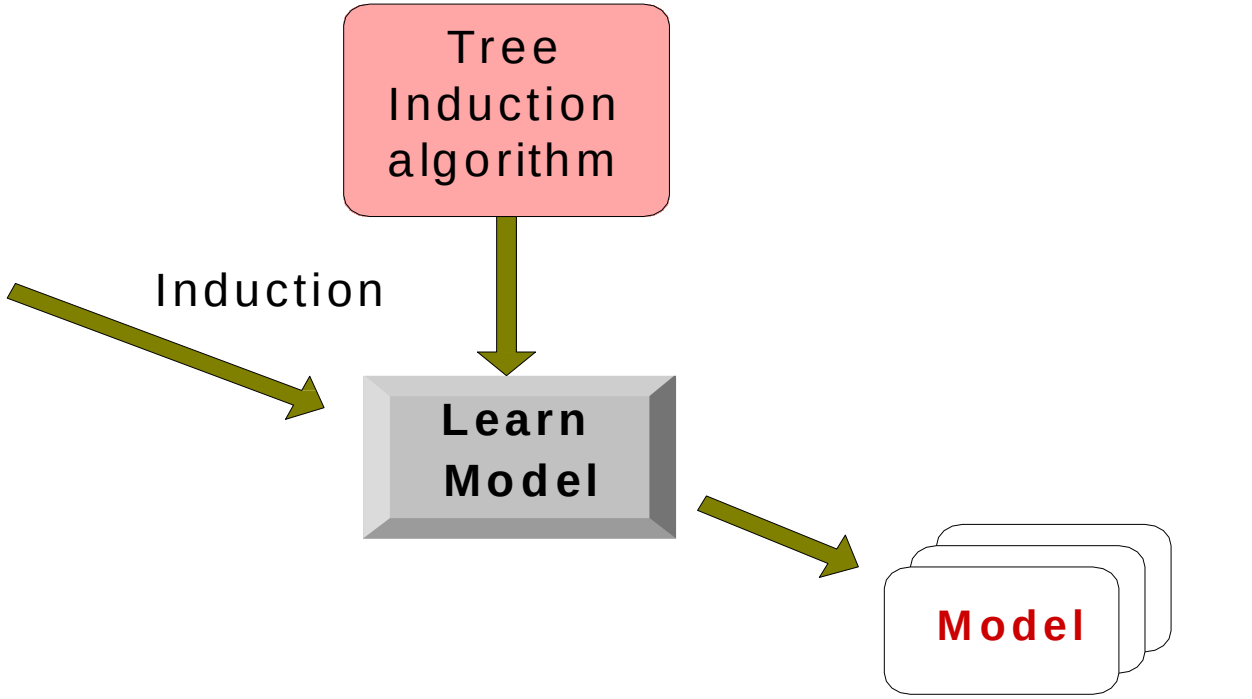
Training Data Matrix

Decision Tree Model

DECISION TREE INDUCTION AND DEDUCTION

Tid	Attrib1	Attrib2	Attrib3	Class
1	Yes	Large	125K	No
2	No	Medium	100K	No
3	No	Small	70K	No
4	Yes	Medium	120K	No
5	No	Large	95K	Yes
6	No	Medium	60K	No
7	Yes	Large	220K	No
8	No	Small	85K	Yes
9	No	Medium	75K	No
10	No	Small	90K	Yes

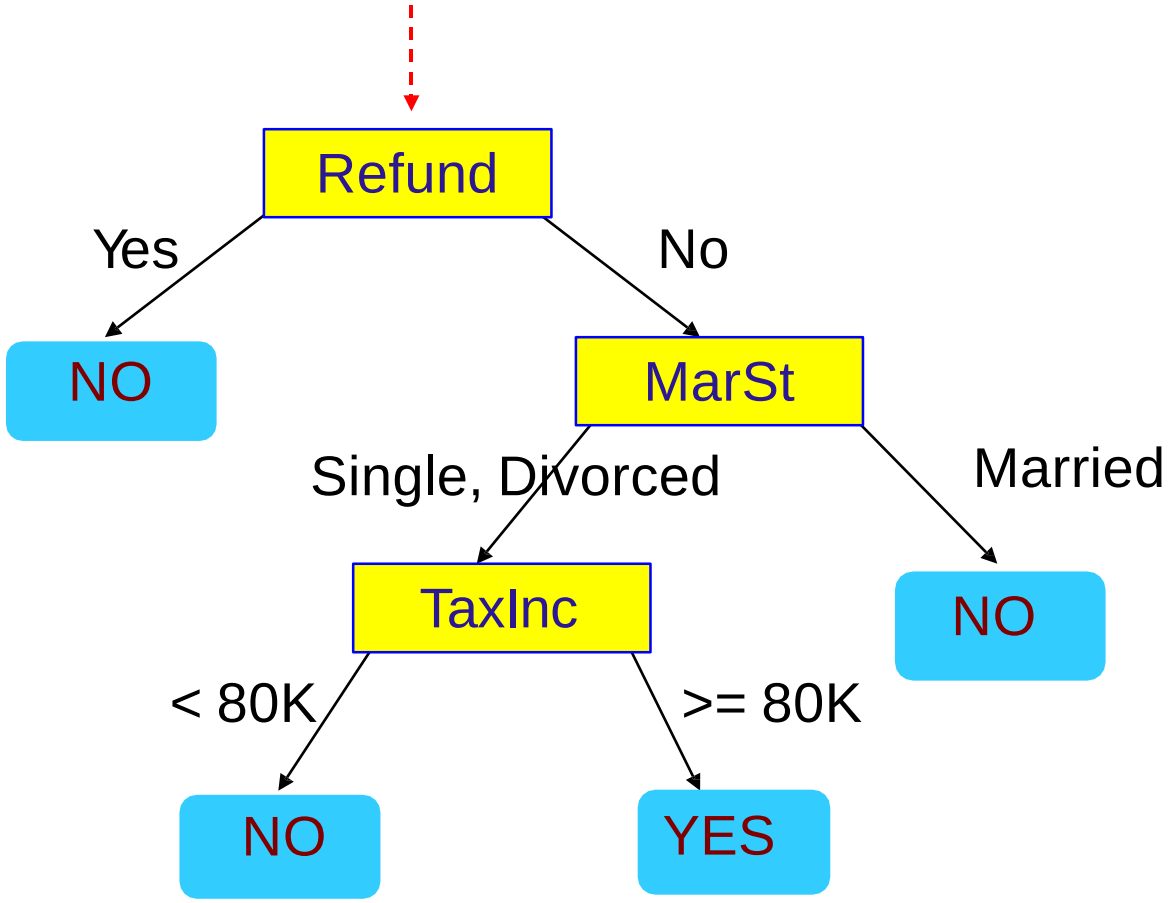
Training Set



Decision Tree

DECISION TREE APPLY MODEL

Start from the root of tree.



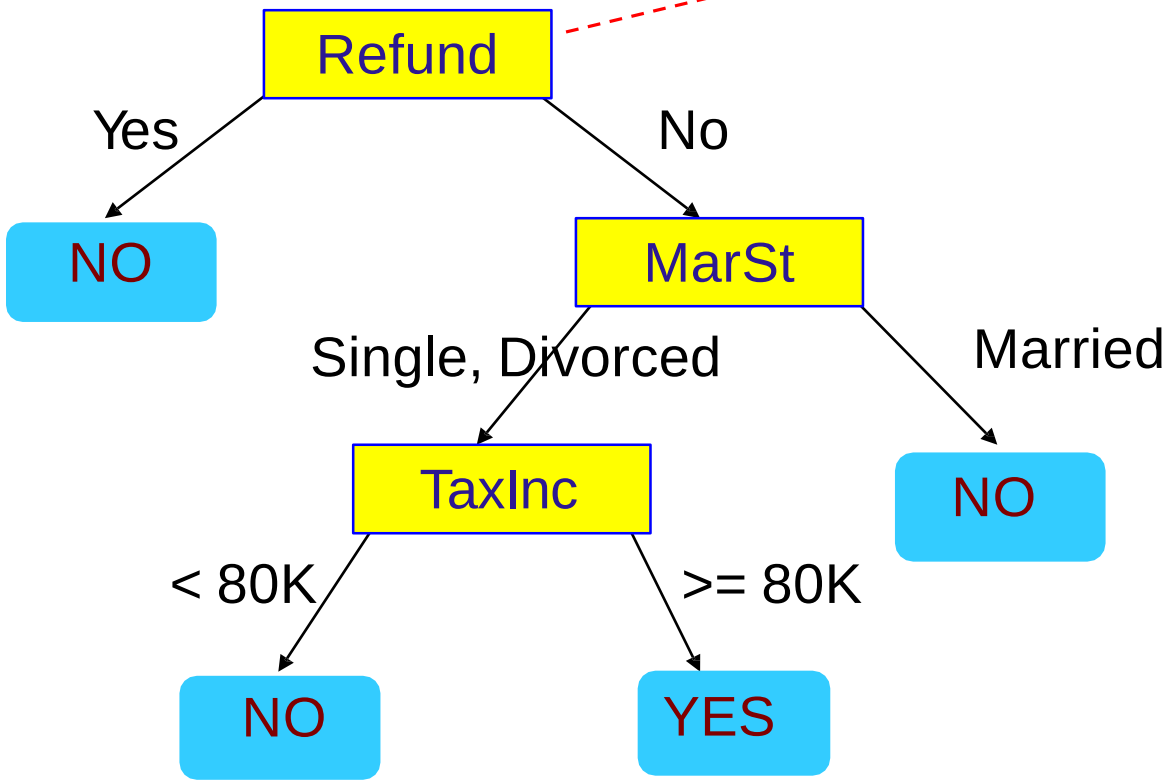
Test Data

Refund	Marital Status	Taxable Income	Fraud
No	Married	80K	?

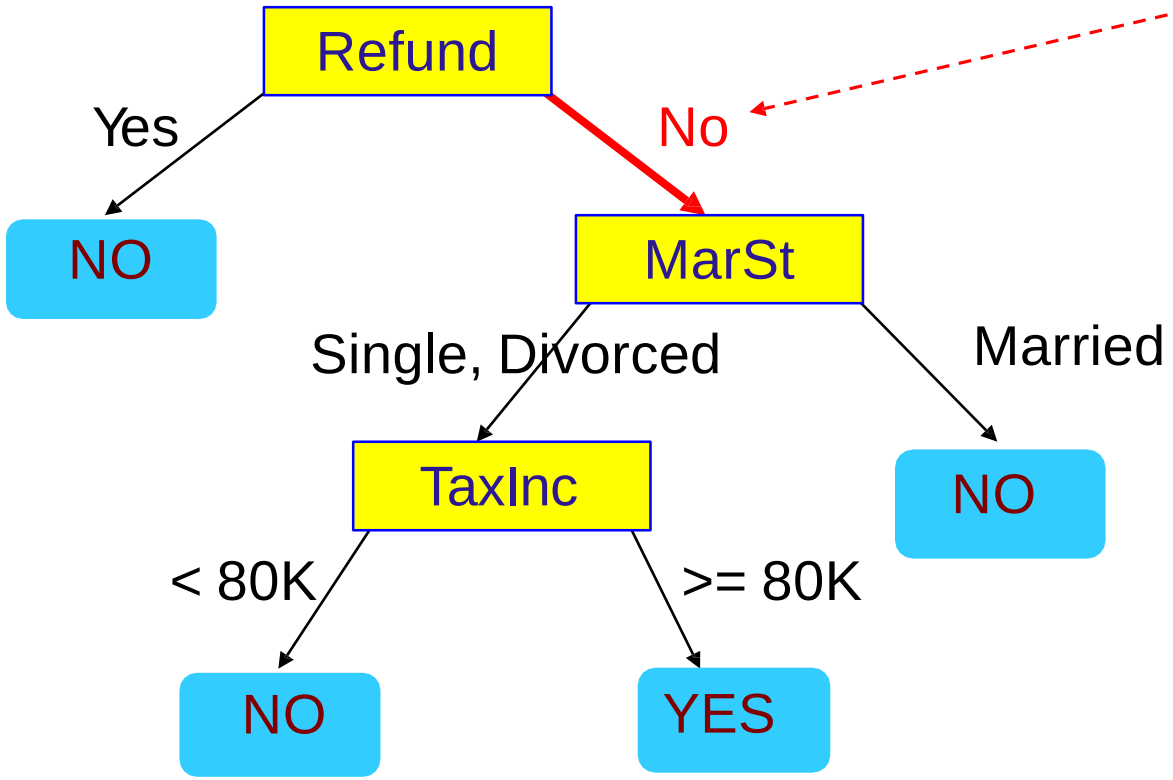
DECISION TREE APPLY MODEL

Test Data

Refund	Marital Status	Taxable Income	Fraud
No	Married	80K	?



DECISION TREE APPLY MODEL



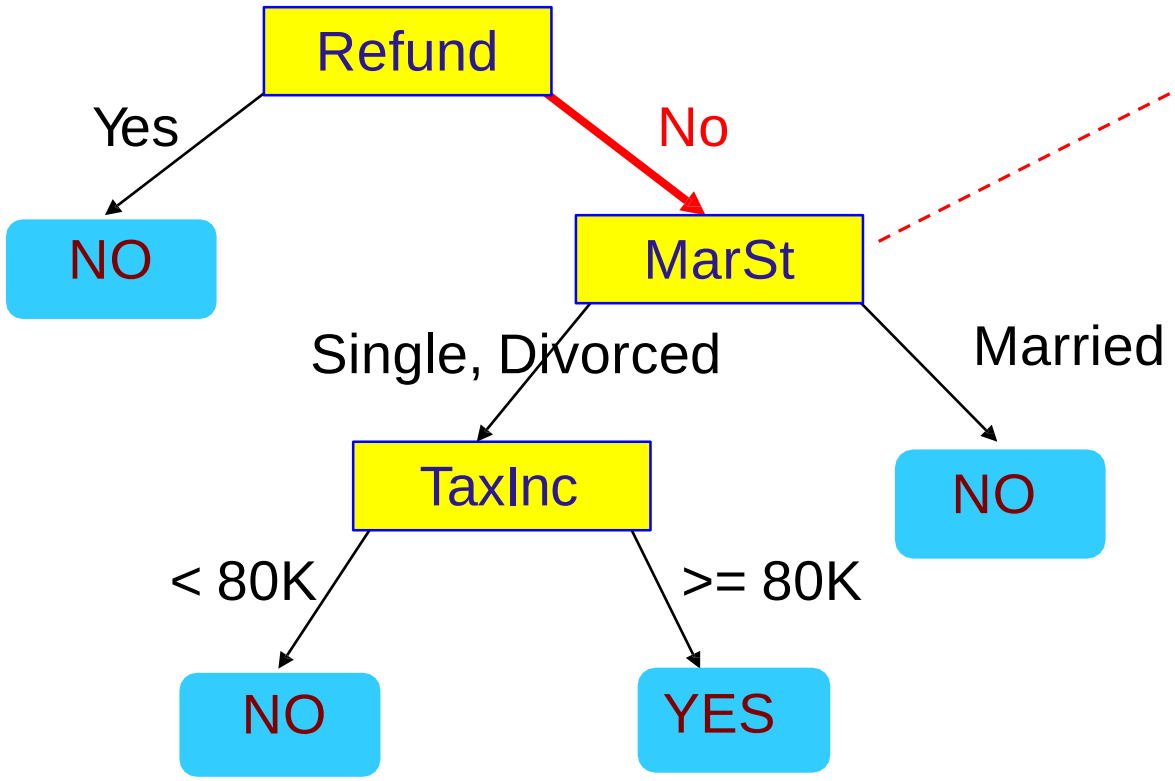
Test Data

Refund	Marital Status	Taxable Income	Fraud
No	Married	80K	?

DECISION TREE APPLY MODEL

Test Data

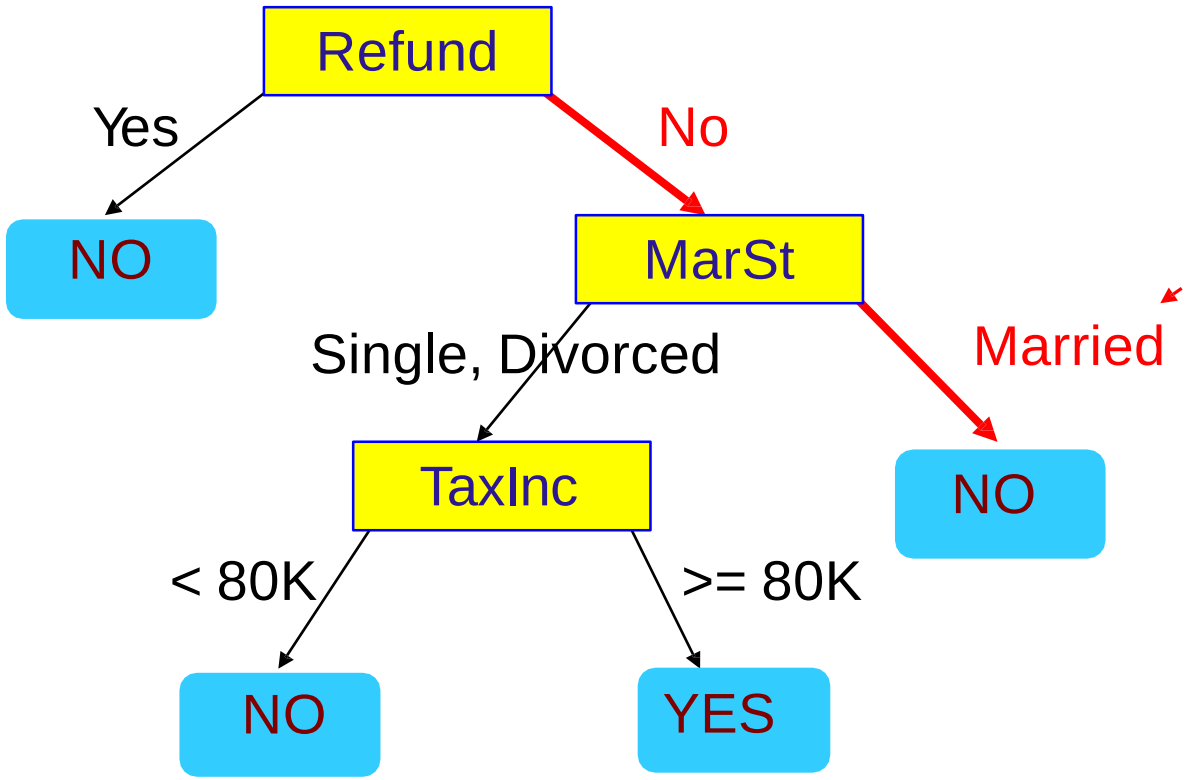
Refund	Marital Status	Taxable Income	Fraud
No	Married	80K	?



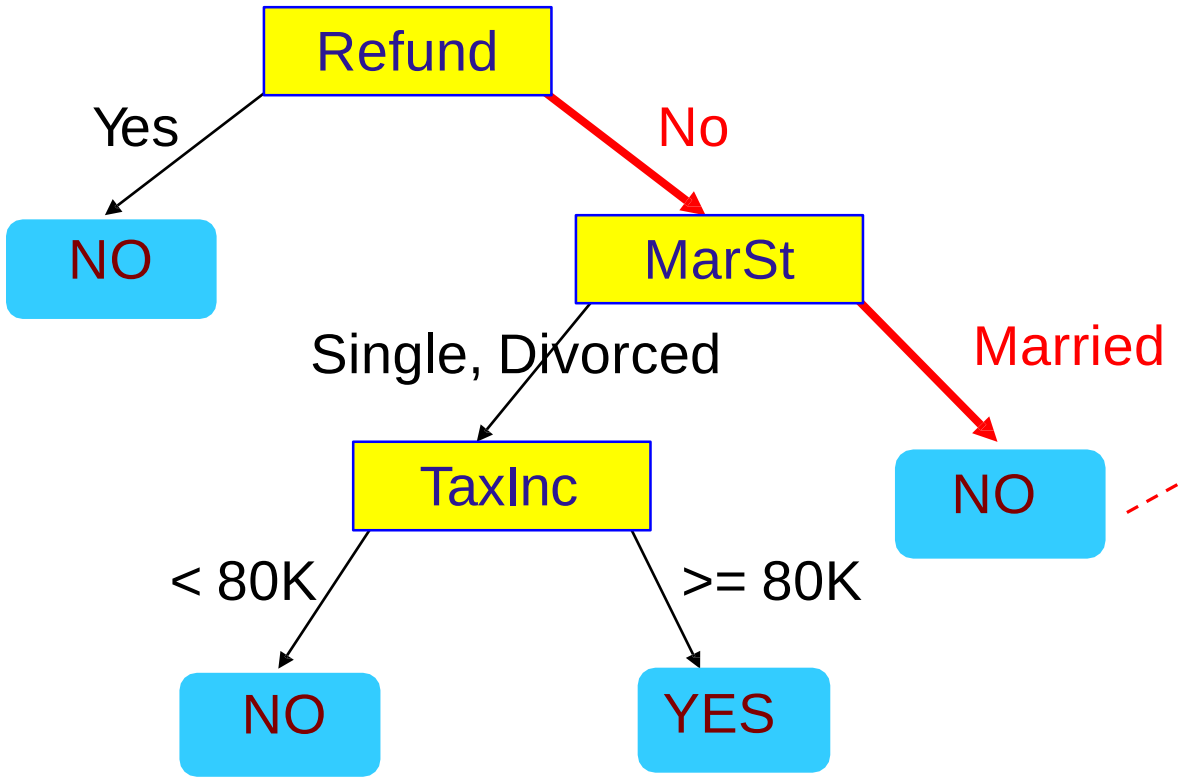
DECISION TREE APPLY MODEL

Test Data

Refund	Marital Status	Taxable Income	Fraud
No	Married	80K	?



DECISION TREE APPLY MODEL



Test Data

Refund	Marital Status	Taxable Income	Fraud
No	Married	80K	?

Assign value of "No" to Fraud