

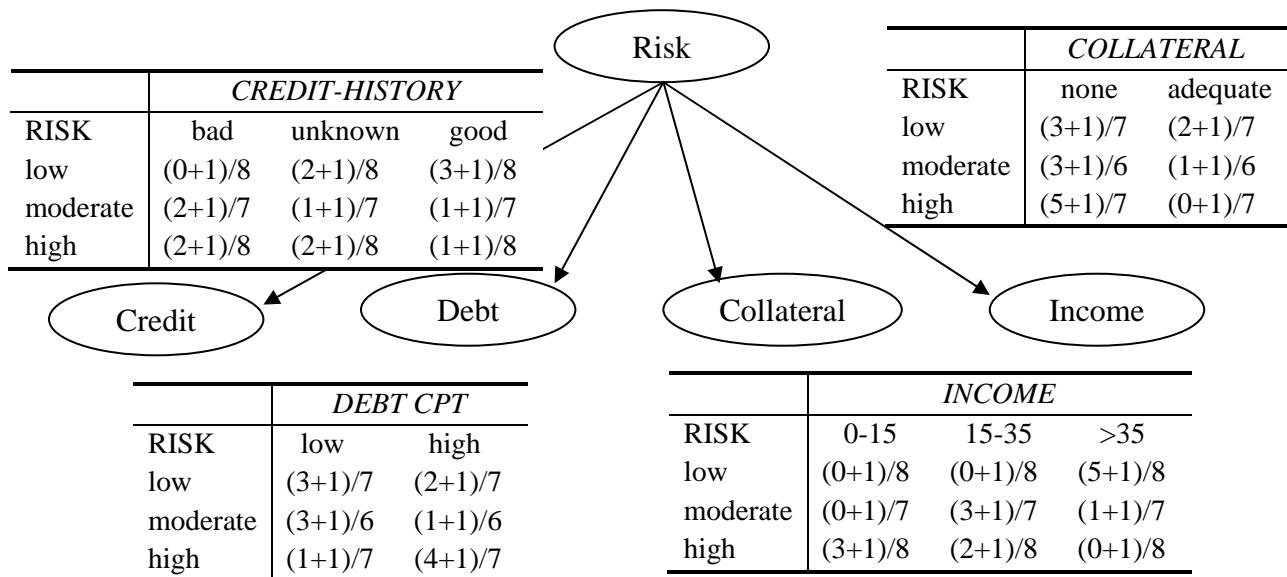
Dataset:

```
@relation credit-data
@attribute credit_history {bad, unknown, good}
@attribute debt {low, high}
@attribute collateral {none, adequate}
@attribute income {0-15, 15-35, >35}
@attribute risk {low, moderate, high}
@data
bad,          low,   none,   0-15,  high
unknown,     high,  none,   15-35, high
unknown,     low,   none,   15-35, moderate
bad,          low,   none,   15-35, moderate
unknown,     low,   adequate, >35,  low
unknown,     low,   none,   >35,  low
unknown,     high,  none,   0-15,  high
bad,          low,   adequate, >35,  moderate
good,         low,   none,   >35,  low
good,         high,  adequate, >35,  low
good,         high,  none,   0-15,  high
good,         high,  none,   15-35, moderate
good,         high,  none,   >35,  low
bad,          high,  none,   15-35, high
```

Naïve Bayes Model Construction:

Constructing a Naïve Bayes model over the above credit-data, where risk is the target attribute.

<i>RISK:</i>	<i>low</i>	<i>moderate</i>	<i>high</i>
probability:	$(5+1)/17$	$(4+1)/17$	$(5+1)/17$



Classification using the Naïve Bayes Model:

Using the above Naïve Bayes model to classify a new instance:

Credit-History	Debt	Collateral	Income	Risk
good	low	adequate	0-15	?

predicted $v = \operatorname{argmax} P(\text{Risk}=v)$

- * $P(\text{Credit-History}=\text{good} \mid \text{Risk}=v)$
- * $P(\text{Debt}=\text{low} \mid \text{Risk}=v)$
- * $P(\text{Collateral}=\text{adequate} \mid \text{Risk}=v)$
- * $P(\text{Income}=0-15 \mid \text{Risk}=v)$

$v = \text{low}: P(\text{Risk}=\text{low})$

- * $P(\text{Credit-History}=\text{good} \mid \text{Risk}=\text{low})$
- * $P(\text{Debt}=\text{low} \mid \text{Risk}=\text{low})$
- * $P(\text{Collateral}=\text{adequate} \mid \text{Risk}=\text{low})$
- * $P(\text{Income}=0-15 \mid \text{Risk}=\text{low})$

$$= (6/17) * (4/8) * (4/7) * (3/7) * (1/8)$$
$$= 288/53312 = 0.0054$$

$v = \text{moderate}: P(\text{Risk}=\text{moderate})$

- * $P(\text{Credit-History}=\text{good} \mid \text{Risk}=\text{moderate})$
- * $P(\text{Debt}=\text{low} \mid \text{Risk}=\text{moderate})$
- * $P(\text{Collateral}=\text{adequate} \mid \text{Risk}=\text{moderate})$
- * $P(\text{Income}=0-15 \mid \text{Risk}=\text{moderate})$

$$= (5/17) * (2/7) * (4/6) * (2/6) * (1/7)$$
$$= 80/29988 = 0.0027$$

$v = \text{high}: P(\text{Risk}=\text{high})$

- * $P(\text{Credit-History}=\text{good} \mid \text{Risk}=\text{high})$
- * $P(\text{Debt}=\text{low} \mid \text{Risk}=\text{high})$
- * $P(\text{Collateral}=\text{adequate} \mid \text{Risk}=\text{high})$
- * $P(\text{Income}=0-15 \mid \text{Risk}=\text{high})$

$$= (6/17) * (2/8) * (2/7) * (1/7) * (4/8)$$
$$= 96/53312 = 0.0018$$

Hence, the predicted value by the Naïve Bayes classifier is Risk=low.
