Machine Learning

Naïve Bayes Model

Naïve Bayes Algorithm

- The naïve Bayes algorithm finds its roots in statistics and probability theory.

The Bayes theorem

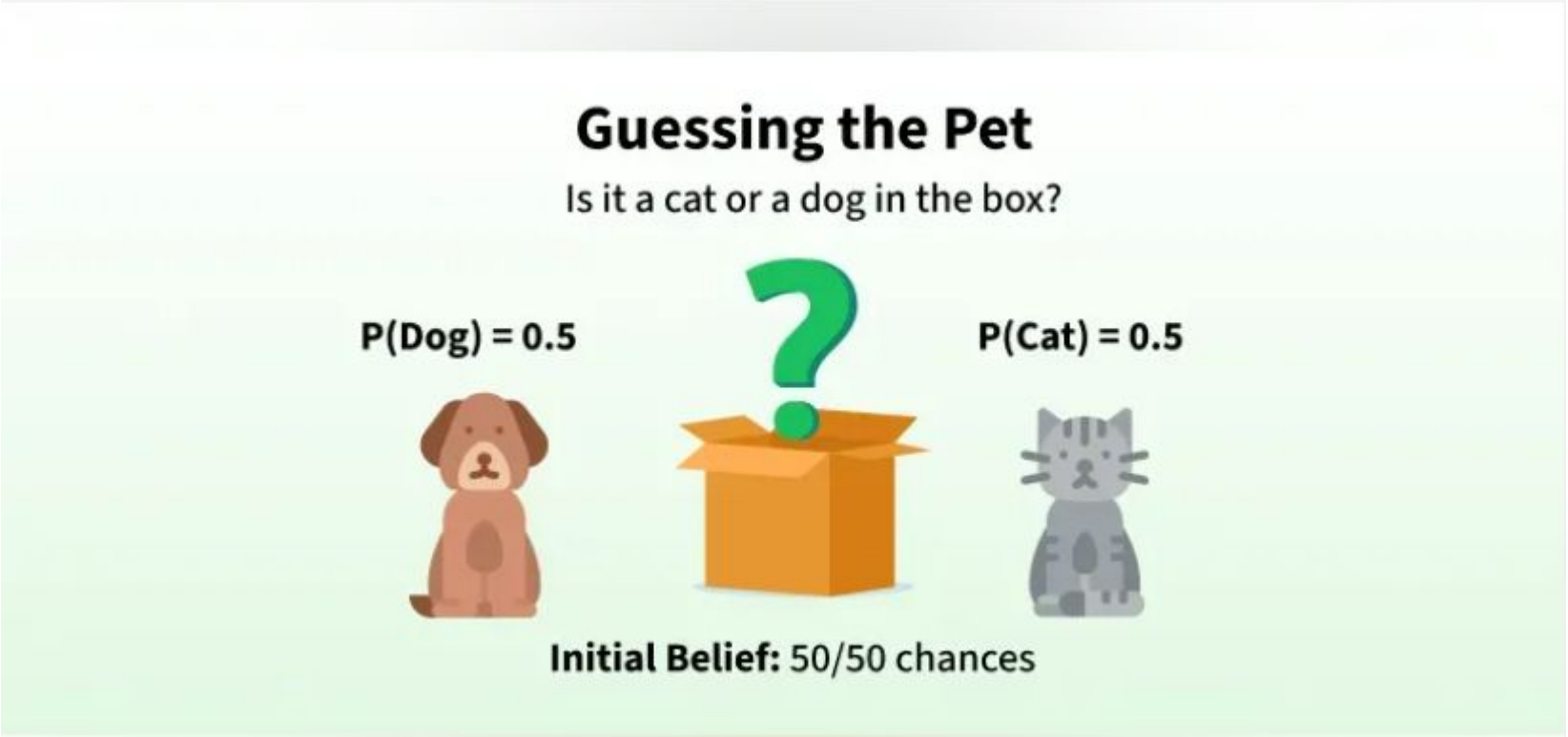
- The naïve Bayesian algorithm is built on the Bayes' theorem, named after Thomas Bayes.
- The Bayes' theorem is one of the most influential and important concepts in statistics and probability theory.

The Bayes theorem

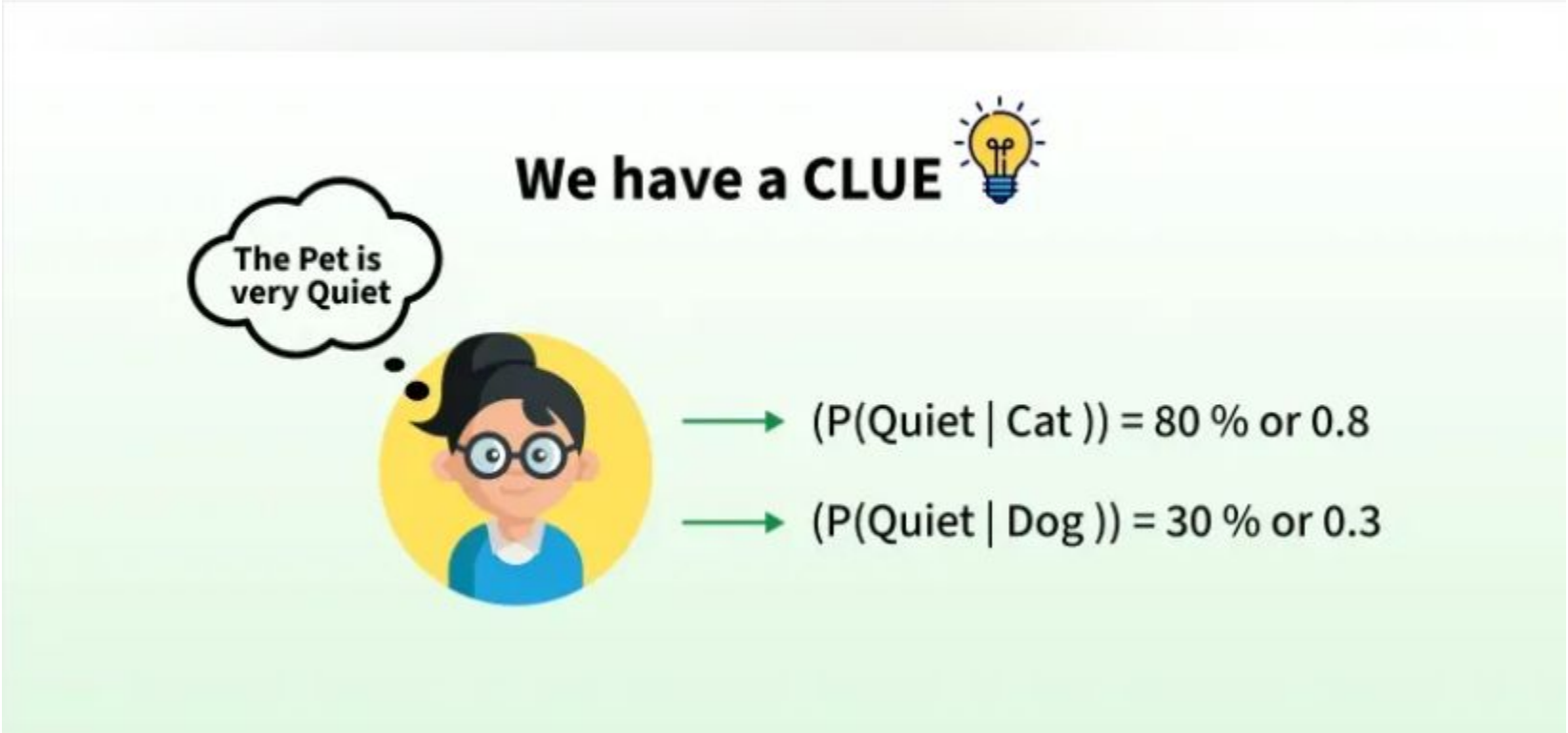
- The Bayes' theorem provides a mathematical expression for
  - how a degree of subjective belief changes to account for new evidence.

The Bayes theorem

The Bayes theorem



The Bayes theorem



The Bayes theorem

Applying Bayes Theorem to Find Probability

$$P(\text{Cat} \mid \text{Quiet}) = \frac{P(\text{Quiet} \mid \text{Cat}) \times P(\text{Cat})}{P(\text{Quiet})} = 72.7\%$$

$$P(\text{Dog} \mid \text{Quiet}) = \frac{P(\text{Quiet} \mid \text{Dog}) \times P(\text{Dog})}{P(\text{Quiet})} = 27.3\%$$

The Naïve Bayes theorem

- Derive the expression for a Naïve Bayes' Classifier from conditional probability.

The Bayes theorem

Table 4.4 Golf Dataset With Modified Temperature and Humidity Attributes

No.	Temperature X_1	Humidity X_2	Outlook X_3	Wind X_4	Play (Class Label) Y
1	High	Med	Sunny	false	no
2	High	High	Sunny	true	no
3	Low	Low	Rain	true	no
4	Med	High	Sunny	false	no
5	Low	Med	Rain	true	no
6	High	Med	Overcast	false	yes
7	Low	High	Rain	false	yes
8	Low	Med	Rain	false	yes
9	Low	Low	Overcast	true	yes
10	Low	Low	Sunny	false	yes
11	Med	Med	Rain	false	yes
12	Med	Low	Sunny	true	yes
13	Med	High	Overcast	true	yes
14	High	Low	Overcast	false	yes

The Naïve Bayes theorem

- Three steps:

Step 1: Calculating Prior Probability

Step 2: Calculating Class Conditional Probability

Step 3: Predicting the Outcome using Bayes' Theorem

The Naïve Bayes theorem

Step 1: Calculating Prior Probability

The Naïve Bayes theorem

Step 2: Calculating Class Conditional Probability

The Naïve Bayes theorem

Step 3: Predicting the Outcome using Bayes' Theorem

Advantage

- Bayesian modeling is quite robust in handling missing values.

Issues

- Incomplete Training Set
- Continuous Attributes
- Attribute Independence