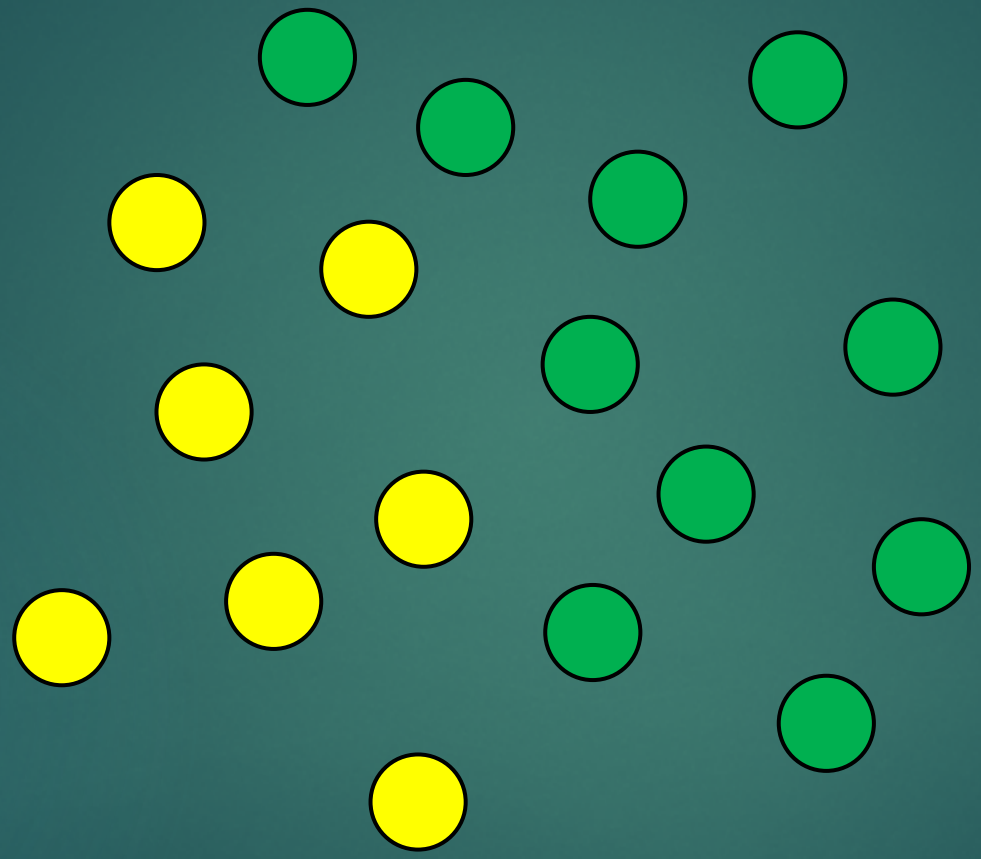
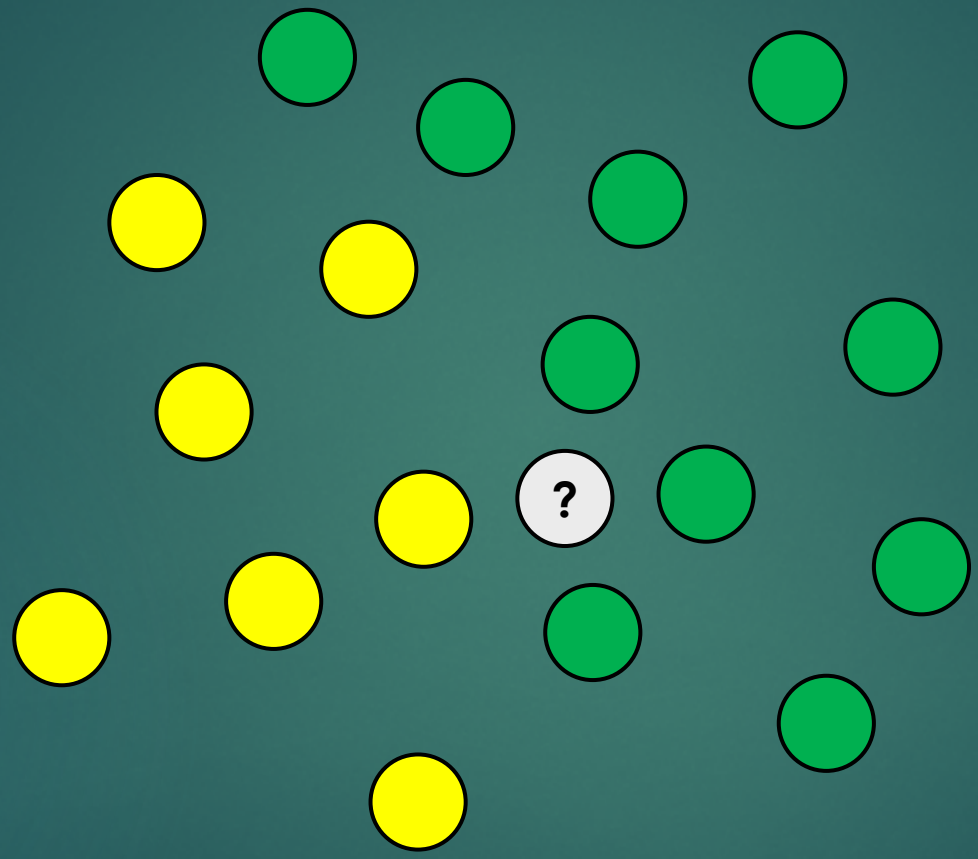
MACHINE LEARNING

NAIVE BAYES

Naive Bayes classifier

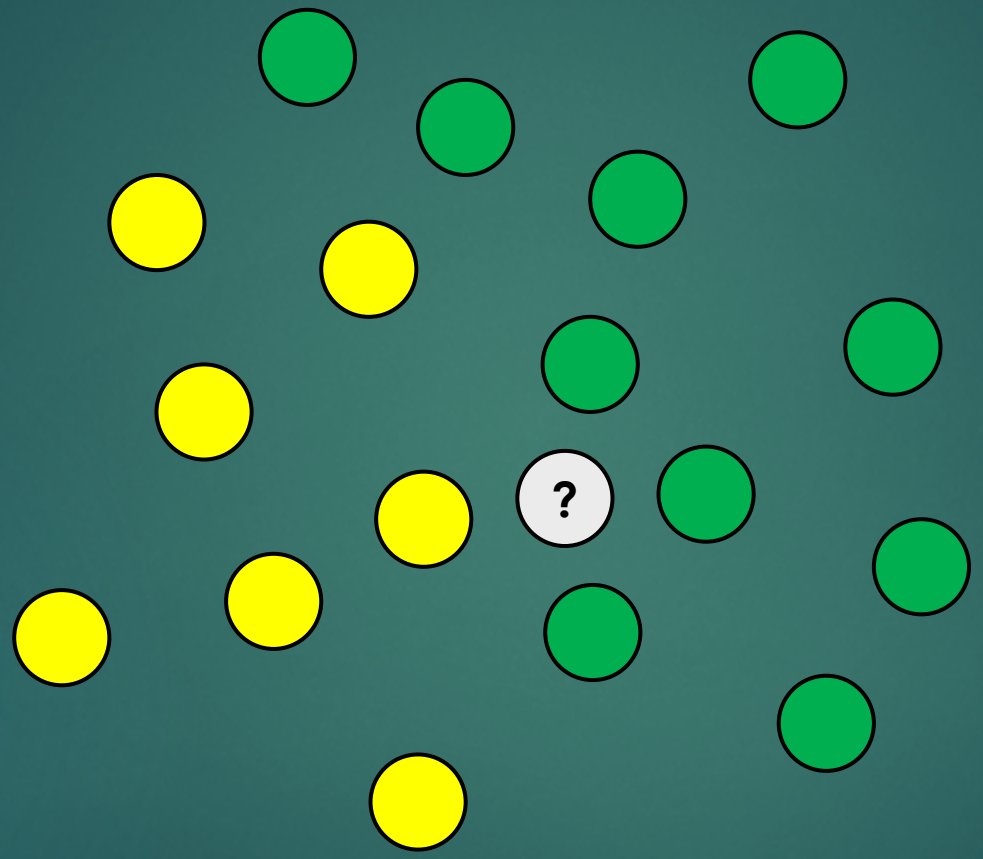
- Very efficient supervised learning algorithm
- It scales well even in high dimensions !!!
- It is able to compete with ***SVM*** or ***random forest*** classifiers
- It is able to make good predictions even when the training data is relatively small
- **Why is it naive?**
- The naive assumption is that every pair of features are independent





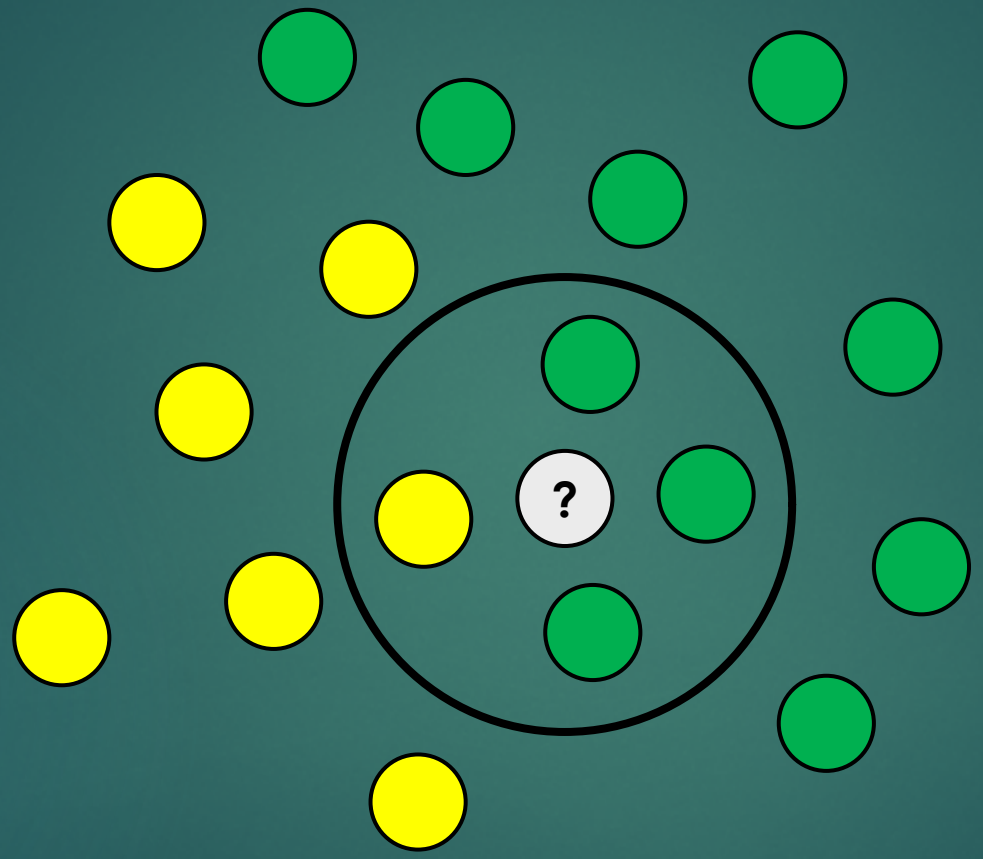
P(yellow) = $\frac{7}{17}$

P(green) = $\frac{10}{17}$



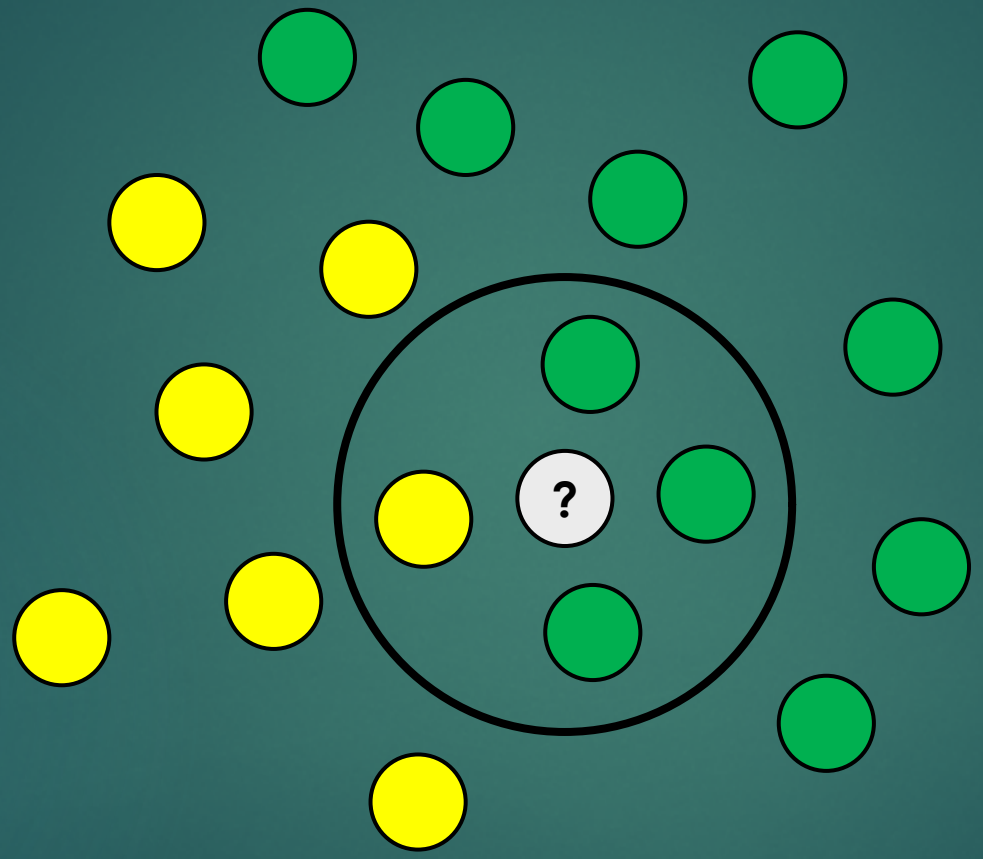
P(yellow) = $\frac{7}{17}$

P(green) = $\frac{10}{17}$



P(yellow) = $\frac{7}{17}$

P(green) = $\frac{10}{17}$

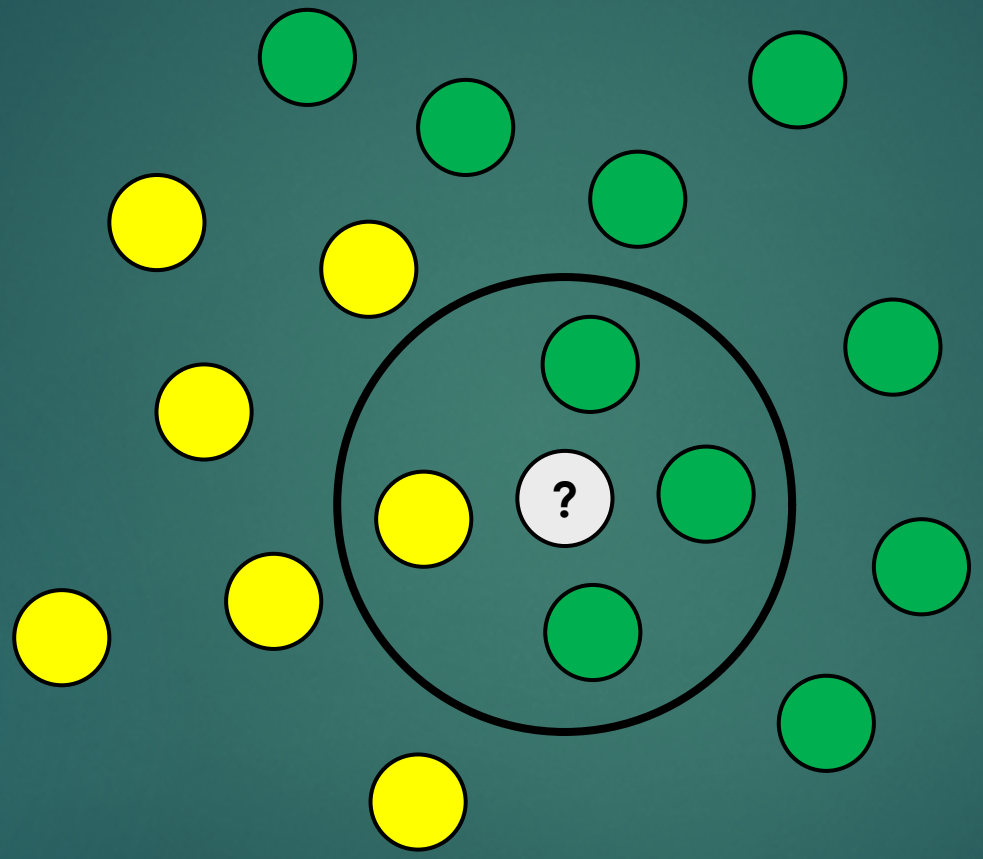


P'(? | green) = $\frac{3}{10}$

P'(? | yellow) = $\frac{1}{7}$

P(yellow) = $\frac{7}{17}$

P(green) = $\frac{10}{17}$



P'(? | green) = $\frac{3}{10}$

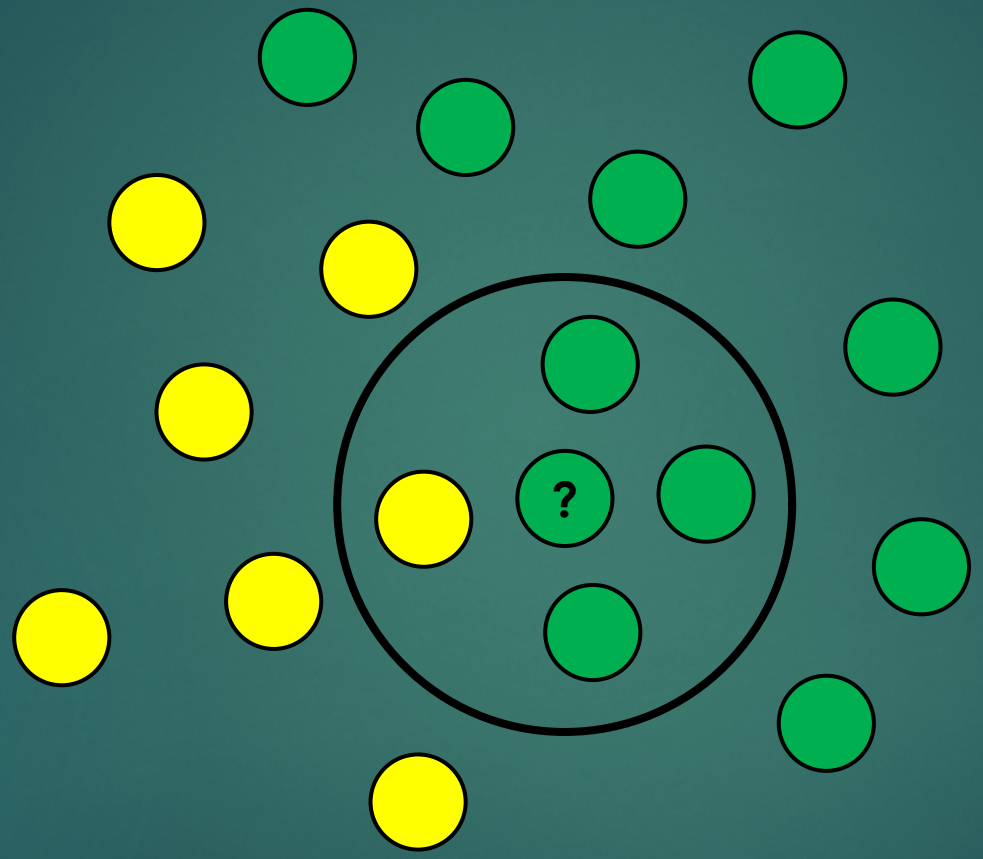
P'(? | yellow) = $\frac{1}{7}$

posterior probability P''(? is green) = P(green) * P'(? | green) = $\frac{10}{17} * \frac{3}{10} = \frac{30}{170}$

posterior probability P''(? is yellow) = P(yellow) * P'(? | yellow) = $\frac{7}{17} * \frac{1}{7} = \frac{7}{119}$

P(yellow) = $\frac{7}{17}$

P(green) = $\frac{10}{17}$



P'(? | green) = $\frac{3}{10}$

P'(? | yellow) = $\frac{1}{7}$

posterior probability P''(? is green) = P(green) * P'(? | green) = $\frac{10}{17} * \frac{3}{10} = \frac{30}{170}$

posterior probability P''(? is yellow) = P(yellow) * P'(? | yellow) = $\frac{7}{17} * \frac{1}{7} = \frac{7}{119}$

Pros and cons

ADVANTAGES

- relatively simple to understand
- it can be trained on small datasets as well
- it is a fast approach
- it is not sensitive to irrelevant features

DISADVANTAGES

- it assumes every feature is independent → it is not always true !!!