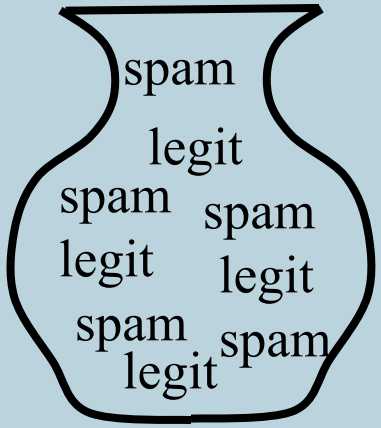
Bayesian / Probabilistic

- Assume data was generated by some sort of structured random process (a so called "generative model").
- Use probability theory to compute the most likely decision (e.g. categorization) assuming the data was generated by this model.
- This computation typically uses "Bayes theorem":

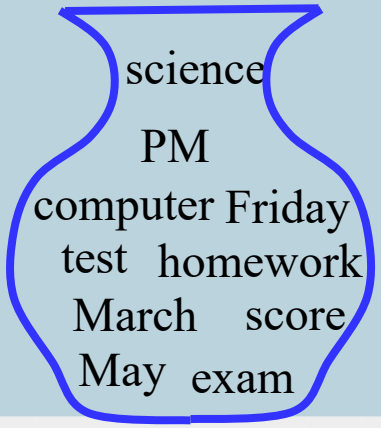
$$P(\text{Hypothesis} \mid \text{Evidence}) = \frac{P(\text{Evidence} \mid \text{Hypothesis})\, P(\text{Hypothesis})}{P(\text{Evidence})}$$

Naïve Bayes Generative Model for Text



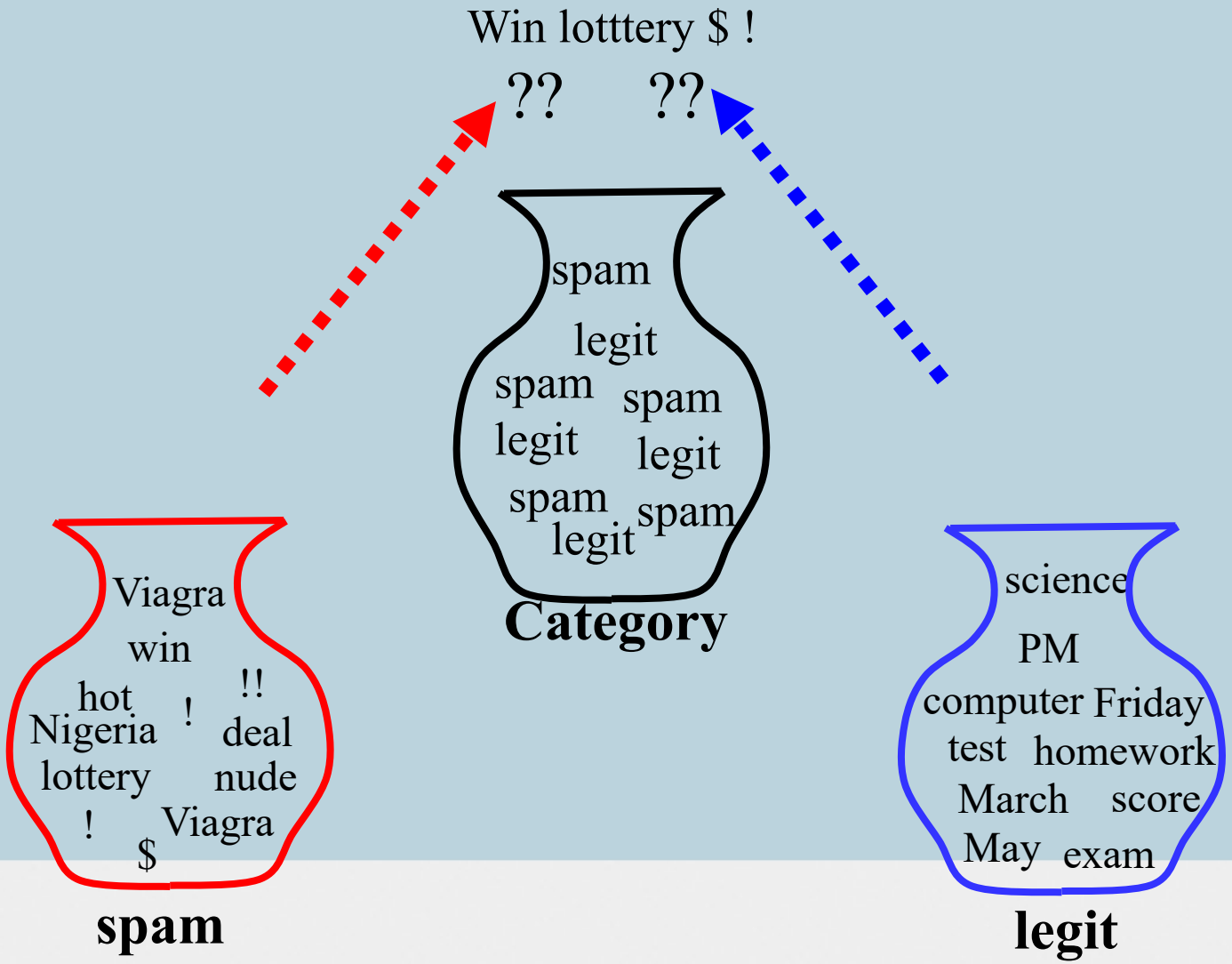
Category

spam



legit

Naïve Bayes Categorization Inference



Learning a Probabilistic Model

- Assume the basic structure of the model or try to first learn it from the data by testing different structures' ability to fit the training data.
- Then estimate the parameters of the model from the training data.
  - Estimate priors: P(category), e.g. P(spam)
  - Estimate conditionals: P(feature | category)

    P(word | spam)

Weaknesses of the Probabilistic Approach

- Assumes generative probabilistic model is (approximately) correct.
  - George Box: "All models are wrong but some are useful"

- Assumes availability of sufficient data to accurately estimate model structure and all parameters.

Connectionist / Neural Network

- Try to computationally model the processing and learning in neurobiological systems / brains.