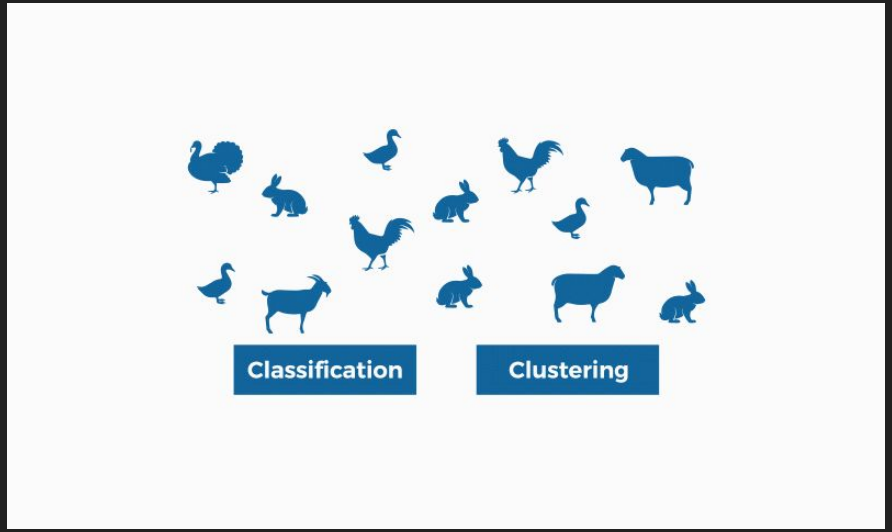
Classification vs. Clustering



Classification
- Predetermined classes
- User defined
-

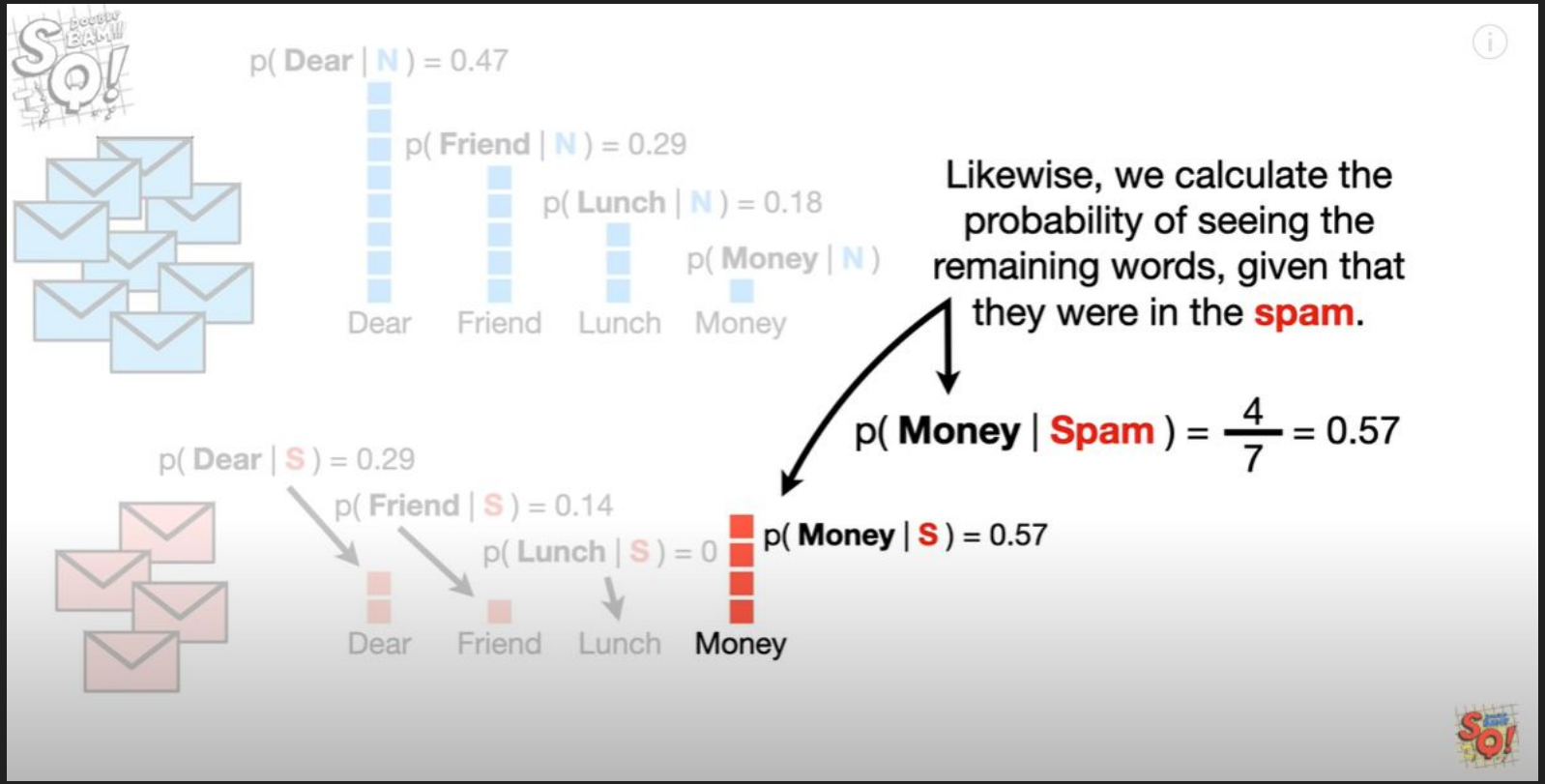
Clustering
- Unsupervised classes
- Machine defined
-

Naive Bayes Classifiers

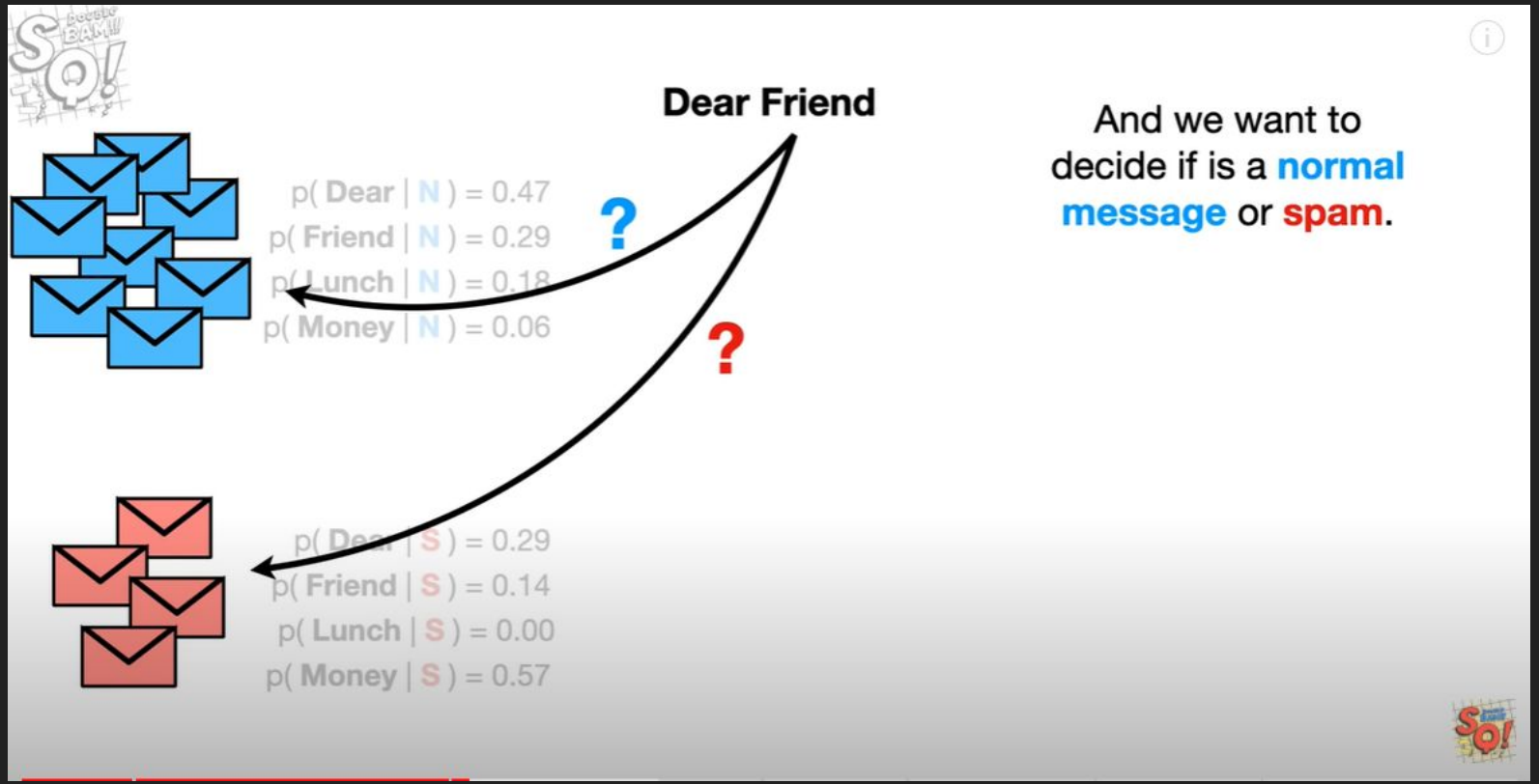
- Classification Technique
- Based on Bayes' Theorem
- Naive Bayes classifiers assume that the presence of a particular feature in a class is unrelated to the presence of any other feature.

- Red
- Round
- 3 Inches in Diameter

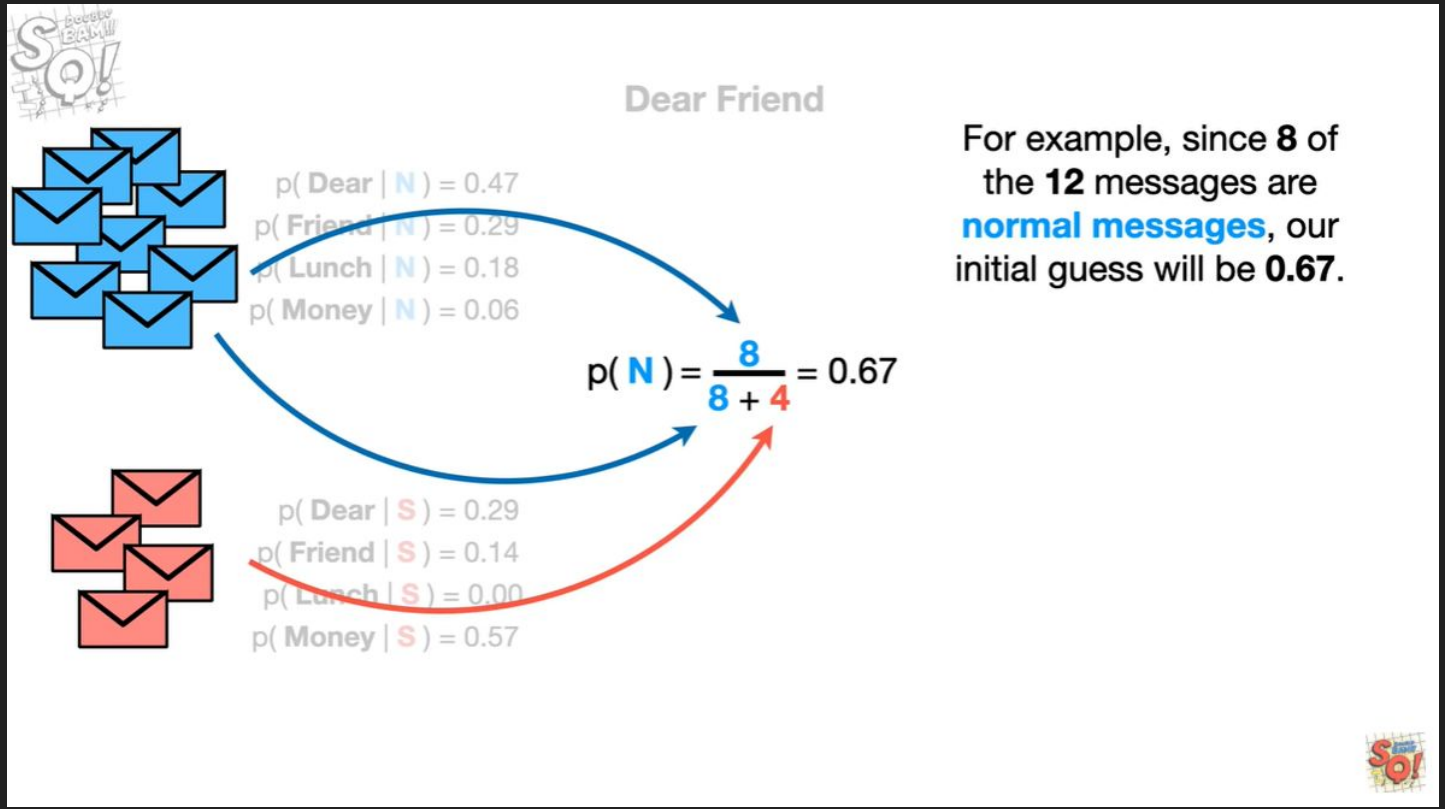
Emails: Naive Bayes Theorem



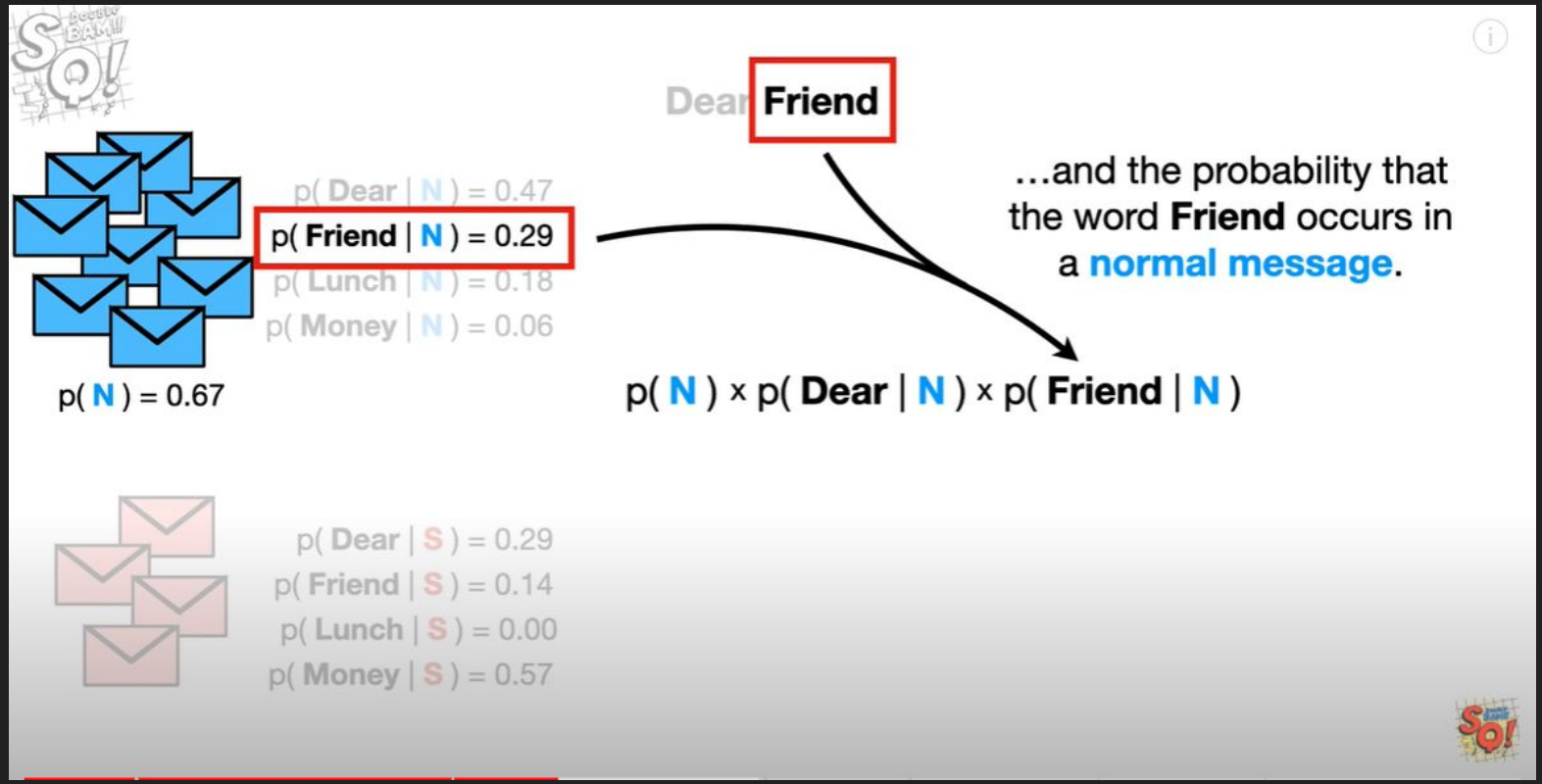
Emails: Naive Bayes Theorem



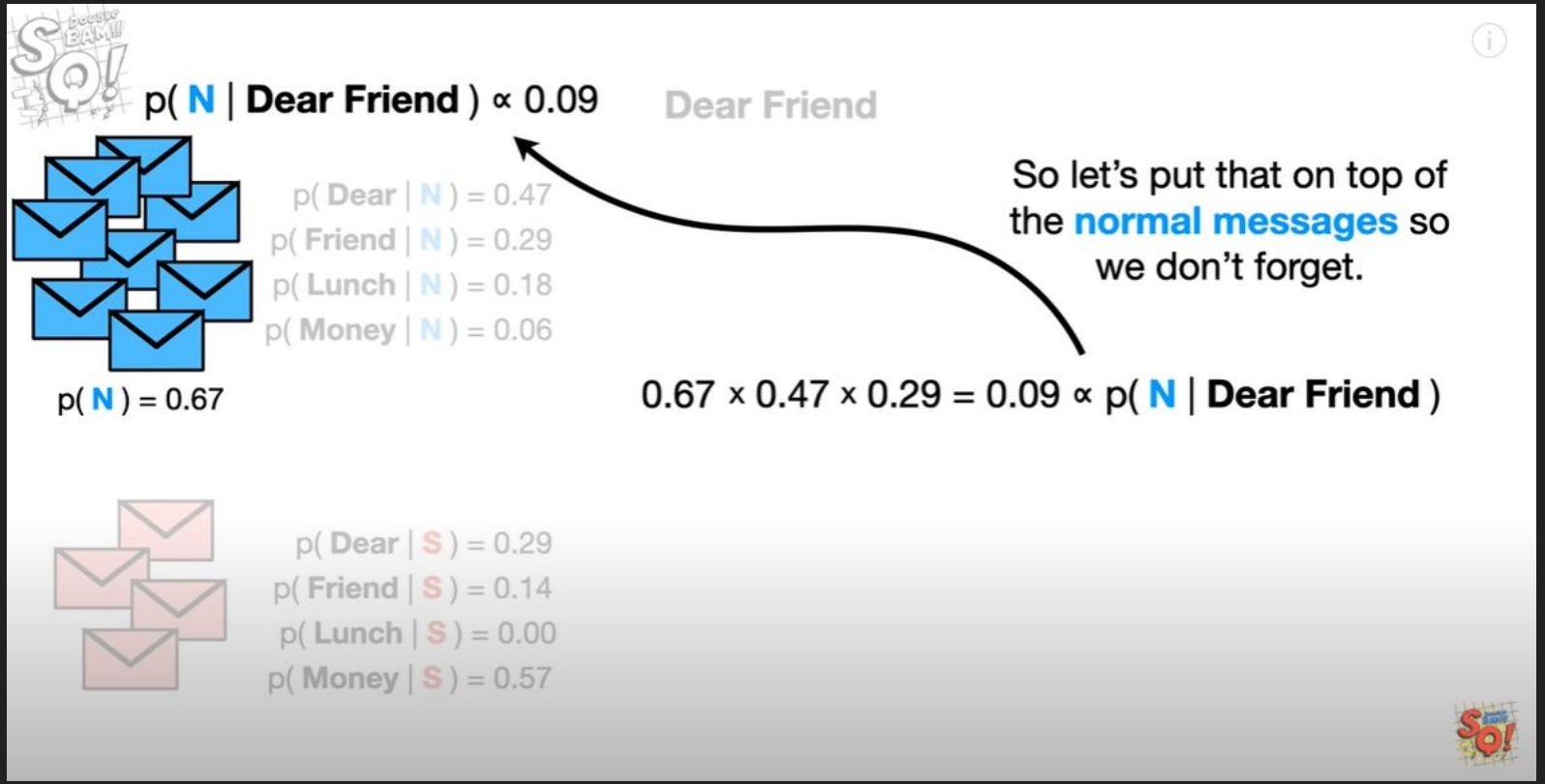
Emails: Naive Bayes Theorem



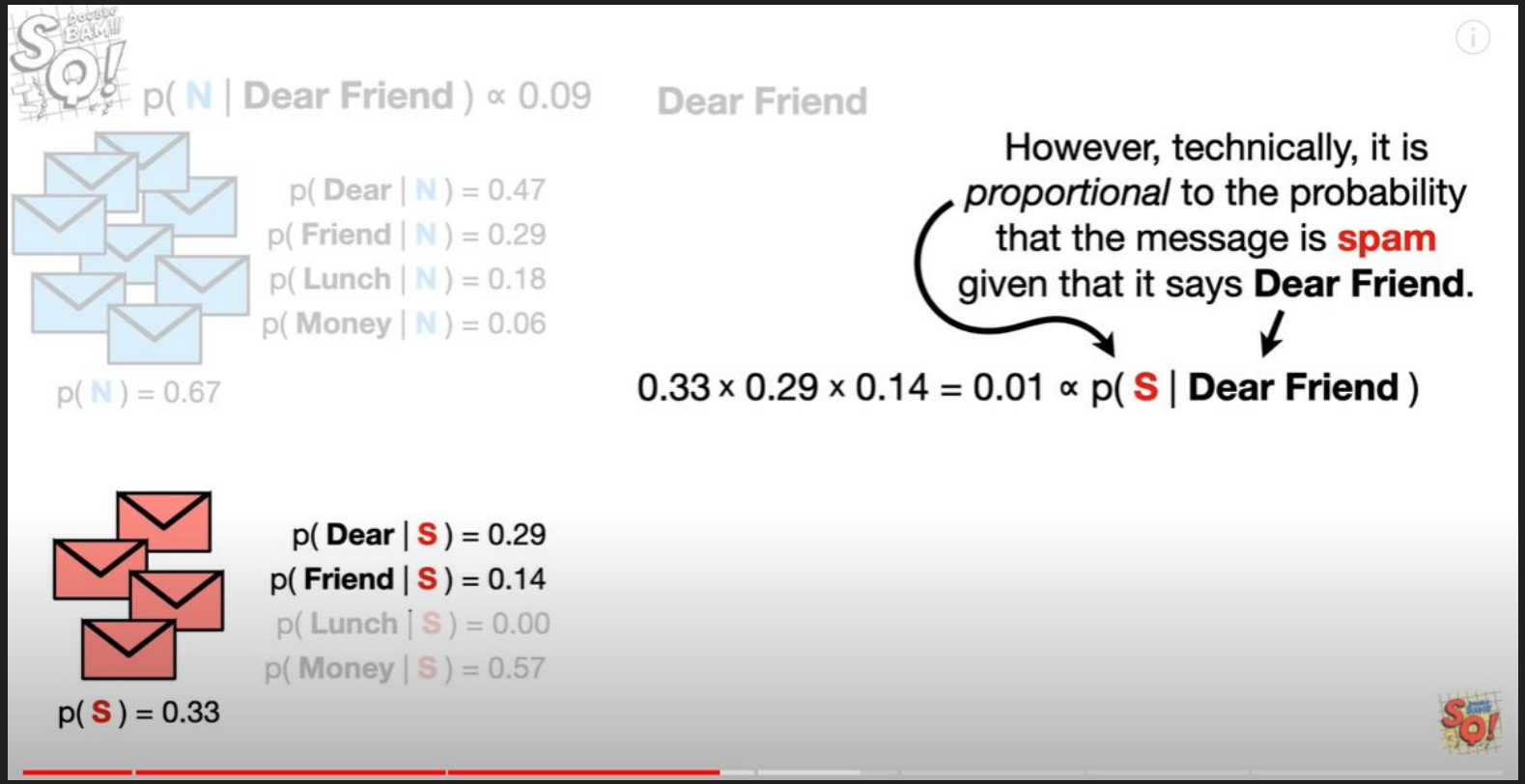
Emails: Naive Bayes Theorem



Emails: Naive Bayes Theorem



Emails: Naive Bayes Theorem



Why so . . . Naive?

- Treats all word orders the same.
- The score for the phrase “Friend Dear” and “Dear Friend” - the exact same.
- Like picking at random from a bag of words.
- But performs very well when ignoring spam
- High Bias
- Low Variance