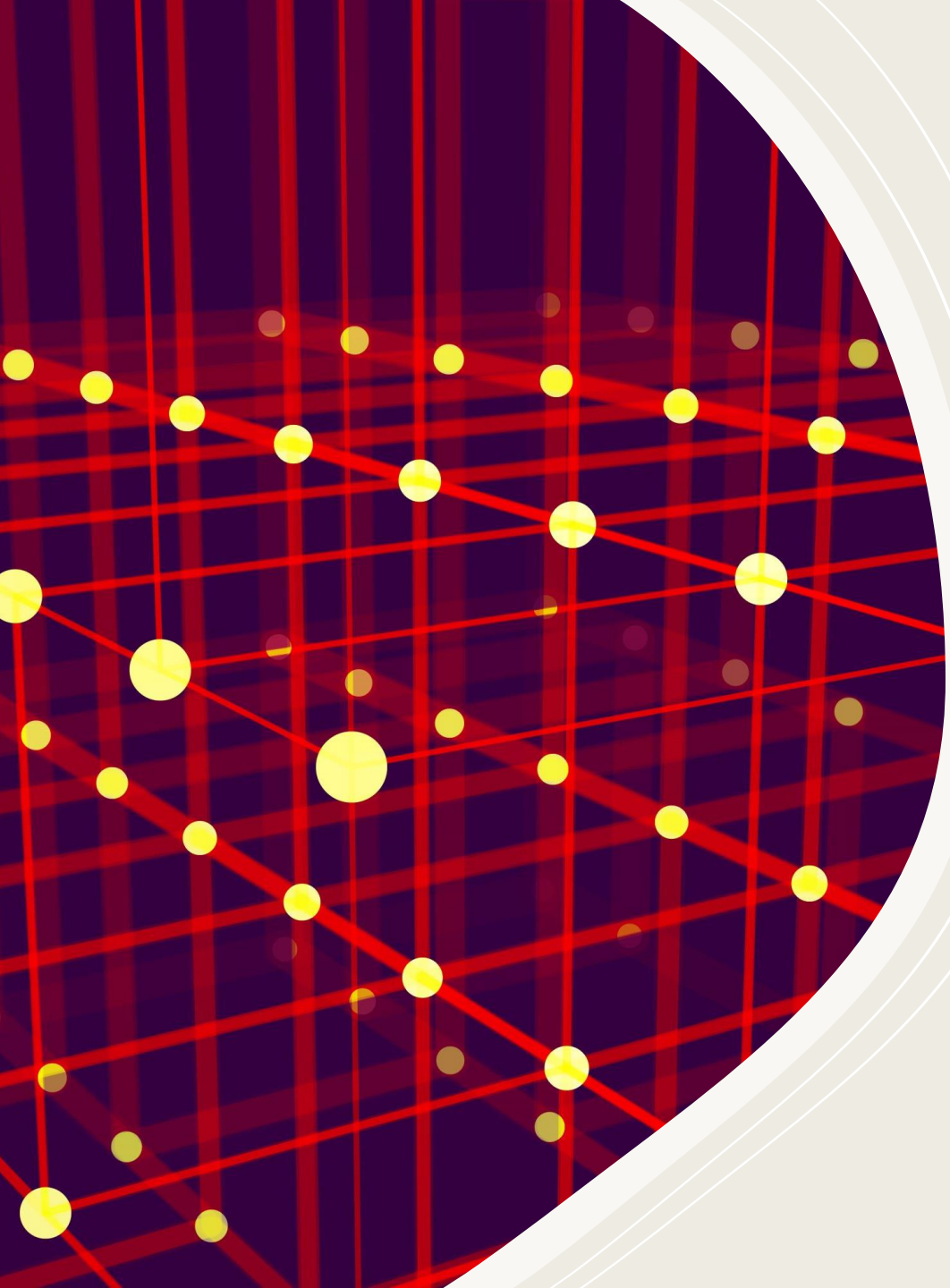


CPSC 340 – Tutorial 2

Lironne Kurzman
lironnek@cs.ubc.ca

University of British Columbia

September 20th, 2021

Training, Testing, and Validation Set

- Given training data, we would like to learn a model to minimize error on the testing data
- How do we decide decision tree depth?
- We care about test error.
- But we can't look at test data.
- So what do we do?????

Training, Testing, and Validation Set

- Given training data, we would like to learn a model to minimize error on the testing data
- One answer: Use part of your train data to approximate test error. Split training objects into training set and validation set:

  Train model on the training data. Test model on the validation data

Cross Validation

- Isn't it wasteful to only use part of your data? k-fold cross-validation:

  Train on k-1 folds of the data, validate on the other fold. Repeat this k times with different splits, and average the score.

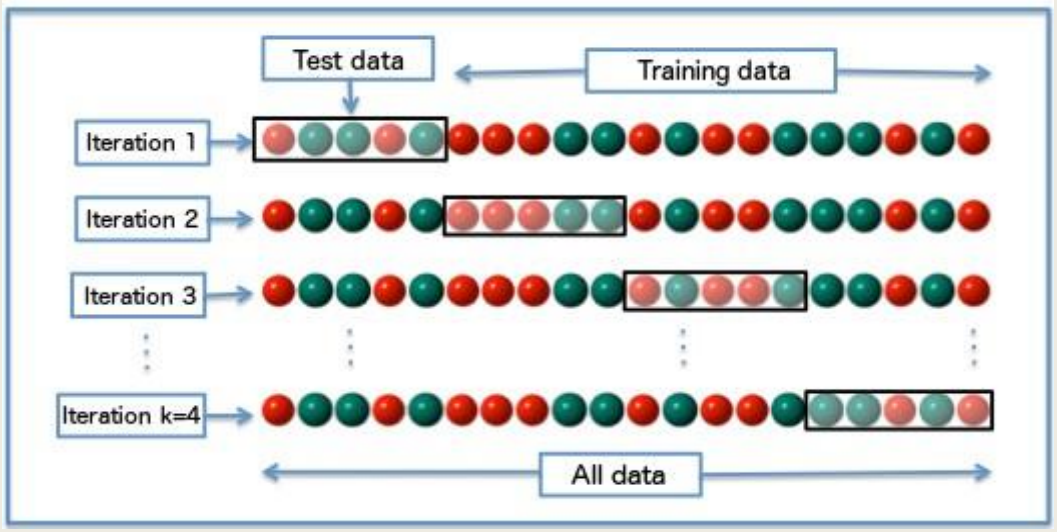
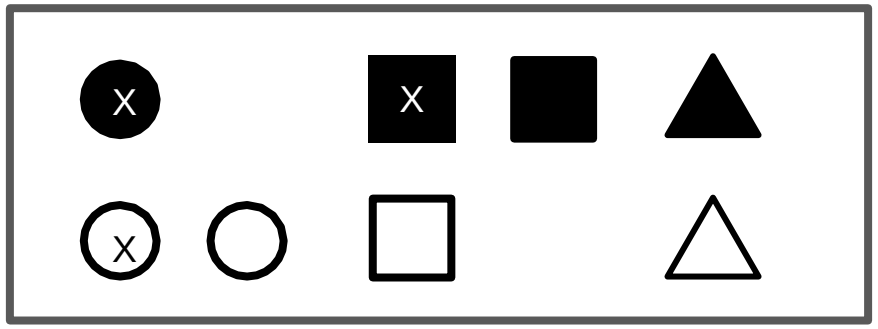


Figure 1: Adapted from Wikipedia

Note: if examples are ordered, split should be random

Naive Bayes: Tips (Frequentist View)

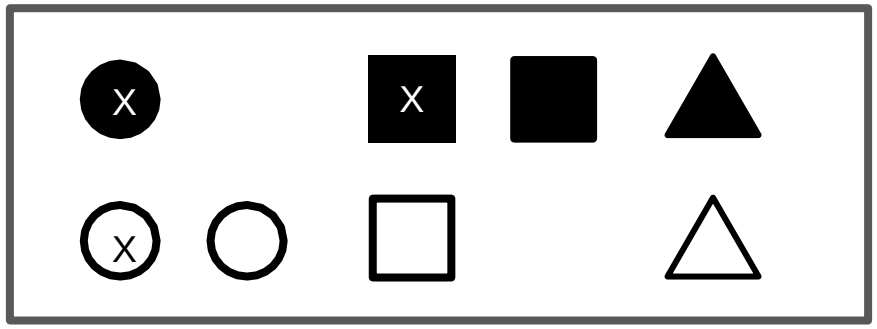


$$P(color = white)$$

:= "**probability** of color being white"

:= "**proportion** of white things"

Naive Bayes: Tips (Frequentist View)

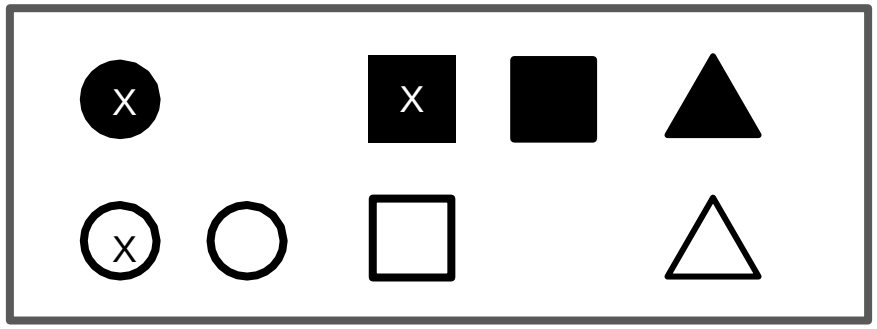


Bayes Theorem:

$$P(Y|X) = \frac{P(X|Y)P(Y)}{P(X)}$$

$$P(\text{ shape} = \text{"circle"} \mid \text{color} = \text{"black"}) = \frac{P(\text{ shape} = \text{"circle"} \ , \ \text{color} = \text{"black"})}{P(\text{color} = \text{"black"})}$$

Naive Bayes: Tips (Frequentist View)

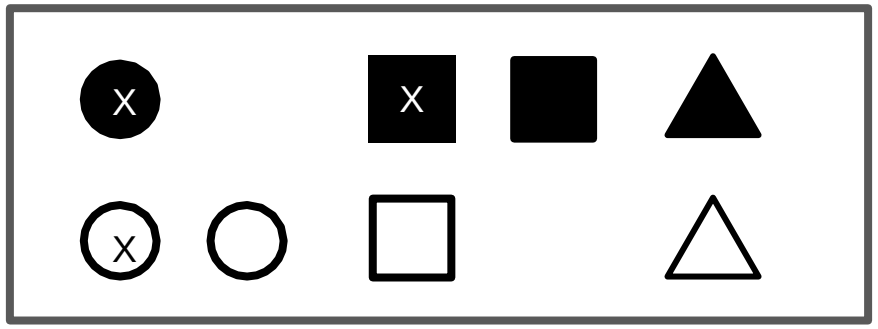


$$P(\text{ shape} = \text{"circle"} \mid \text{color} = \text{"black"}) = \frac{P(\text{ shape} = \text{"circle"} \ , \ \text{color} = \text{"black"})}{P(\text{color} = \text{"black"})}$$

:= "**probability** of shape being circle given color is black"

:= "**proportion** of circles among black items"

Naive Bayes: The Naïve Part



$$P(y_i = marked \mid x_i = [black, square])$$
$$\propto P(x_i = [black, square] \mid y_i = marked) P(y_i = marked)$$
$$\approx P(x_{i1} = black \mid y_i = marked) P(x_{i2} = square \mid y_i = marked) P(y_i = marked)$$

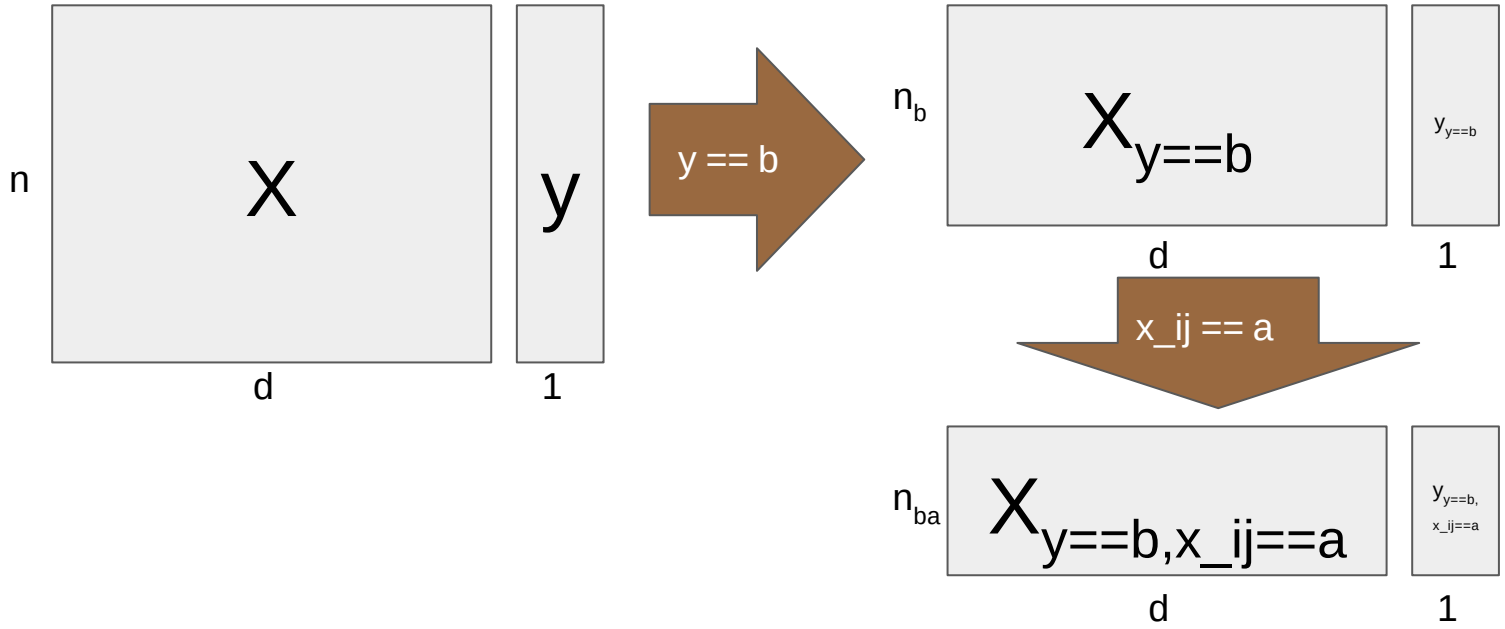
These are our **parameters**.

Naive Bayes in Matrix Form

- Let each feature of X have values {1, 2, 3, …, q}
- Let y have values {1, 2, 3, …, k}

Compute:

$$P(x_{ij} = a \mid y = b)$$

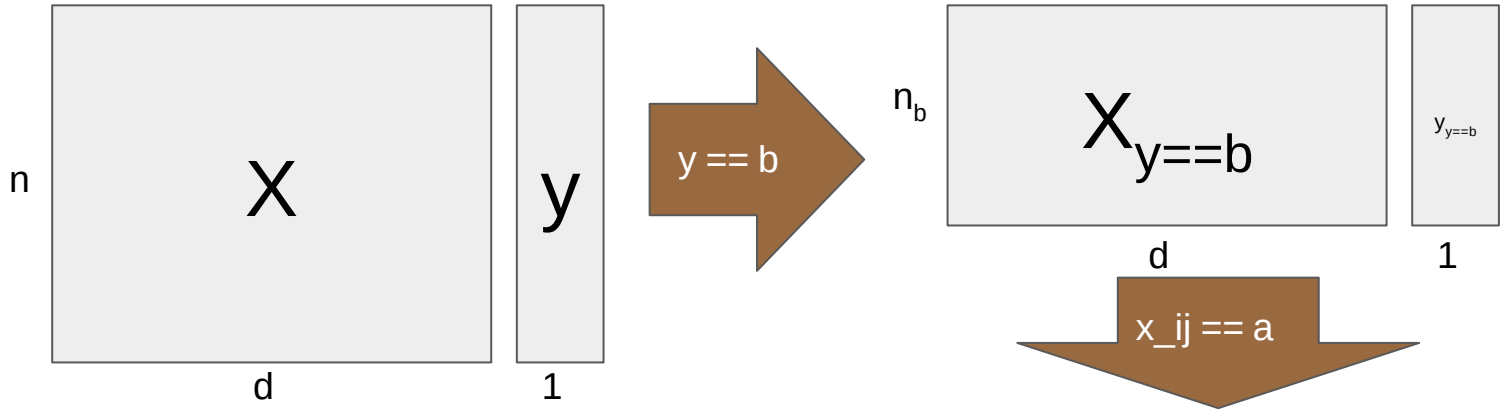


Naive Bayes in Matrix Form

- Let each feature of X have values {1, 2, 3, …, q}
- Let y have values {1, 2, 3, …, k}

Compute:

$$P(x_{ij} = a \mid y = b)$$



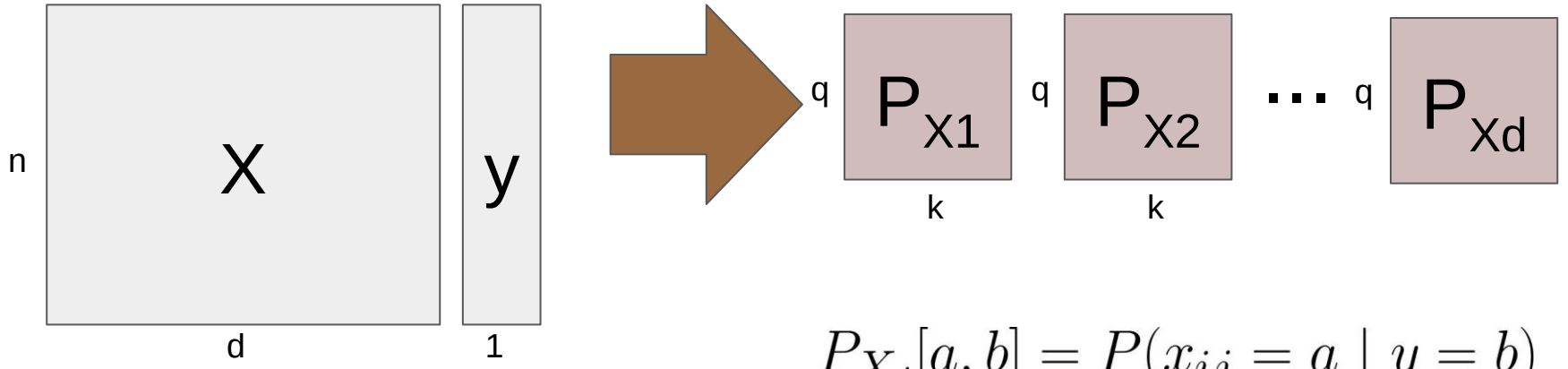
Compute:

$$P(x_{ij} = a \mid y = b)$$

= n_ba / n_b

Naive Bayes in Matrix Form

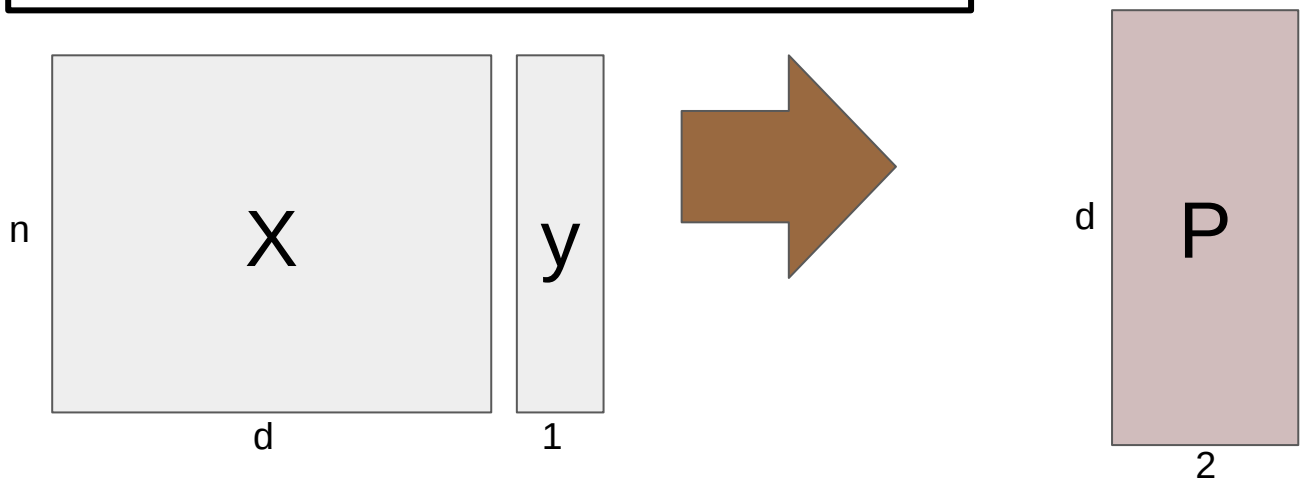
- Let each feature of X have values {1, 2, 3, …, q}
- Let y have values {1, 2, 3, …, k}



$$P_{X_j}[a, b] = P(x_{ij} = a \mid y = b)$$

Naive Bayes in Matrix Form

- Let each feature of X have values {1, 2, 3, …, q}
- Let y have values {1, 2, 3, …, k}



Q: What do the entries of P correspond to?

Naive Bayes Example

headache	*runny nose*	*fever*	*flu*
N	*Y*	*Y*	*N*
Y	*N*	*N*	*N*
N	*N*	*N*	*N*
Y	*Y*	*Y*	*Y*
Y	*Y*	*N*	*Y*
N	*N*	*Y*	*Y*

headache	*runny nose*	*fever*	*flu*
Y	*N*	*Y*	*?*

Naive Bayes Example

$p(headache=Y \mid flu=N)$	1/3
$p(headache=Y \mid flu=Y)$	2/3
$p(runny\ nose=N \mid flu=N)$	2/3
$p(runny\ nose=N \mid flu=Y)$	1/3
$p(fever=Y \mid flu=N)$	1/3
$p(fever=Y \mid flu=Y)$	2/3
$p(flu=N)$	1/2
$p(flu=Y)$	1/2

References

- Based off slides from 2017 and from 2021S by Nam Hee Gordon kim