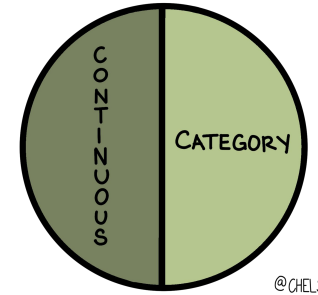


PREDICT



@CHELSEA PARLETT

Naive Bayes and KNN

Dr. Chelsea Parlett-Pelleriti

Naive Bayes

Bayes

$$p(A|B) = \frac{p(B|A) * p(A)}{p(B)}$$

Bayes

$$p(A|B) = \frac{p(B|A) * p(A)}{p(B)}$$

$$p(\text{category} | x_1, x_2, x_3) = \frac{p(x_1, x_2, x_3 | \text{category}) * p(\text{category})}{p(x_1, x_2, x_3)}$$

Bayes

$$p(A|B) = \frac{p(B|A) * p(A)}{p(B)}$$

$$p(\text{heart attack} | S, D, O) = \frac{p(S, D, O | \text{heart attack}) * p(\text{heart attack})}{p(S, D, O)}$$

$$p(\text{no heart attack} | S, D, O) = \frac{p(S, D, O | \text{no heart attack}) * p(\text{no heart attack})}{p(S, D, O)}$$

Bayes

$$p(A|B) = \frac{p(B|A) * p(A)}{p(B)}$$

$$p(\text{heart attack} | S, D, O) = \frac{p(S, D, O | \text{heart attack}) * p(\text{heart attack})}{\cancel{p(S, D, O)}}$$

$$p(\text{no heart attack} | S, D, O) = \frac{p(S, D, O | \text{no heart attack}) * p(\text{no heart attack})}{\cancel{p(S, D, O)}}$$

Bayes

$$p(A|B) = \frac{p(B|A) * p(A)}{p(B)}$$

$$p(\text{heart attack} | S, D, O) \propto p(S, D, O | \text{heart attack}) * p(\text{heart attack})$$

$$p(\text{no heart attack} | S, D, O) \propto p(S, D, O | \text{no heart attack}) * p(\text{no heart attack})$$

Bayes

$$p(A|B) = \frac{p(B|A) * p(A)}{p(B)}$$

$$p(\text{heart attack} | S, D, O) \propto p(S, D, O | \text{heart attack}) * p(\text{heart attack})$$

$$p(\text{no heart attack} | S, D, O) \propto p(S, D, O | \text{no heart attack}) * p(\text{no heart attack})$$

Naive

$$p(S, D, O) = P(S|D, O) * P(D|O) * P(O)$$

$$p(S, D, O) = P(S) * P(D) * P(O)$$

Naive

$$p(S, D, O) = P(S|D, O) * P(D|O) * P(O)$$

$$p(S, D, O) = P(S) * P(D) * P(O)$$

Naive

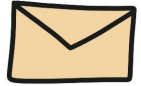
392	business
1	0
0	0
0	0
0	1
0	1
1	1
1	1
1	1
1	1
0	0
0	0
1	1
1	1
0	0
1	0
0	0
1	0
0	1
0	0
1	0

$$p(392|business) \neq p(392)$$

Example



ham

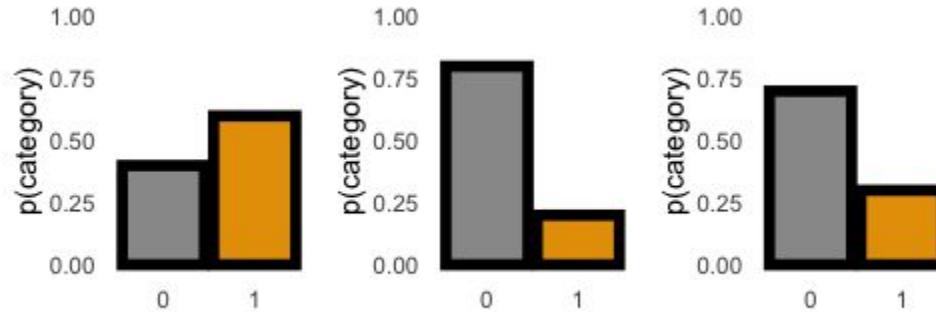


spam

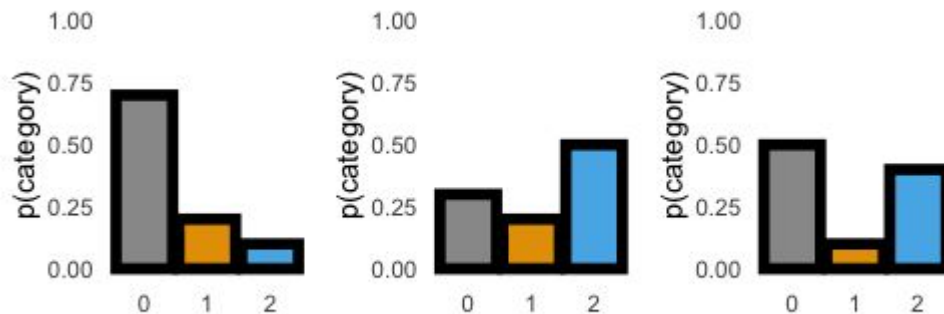
spam	viagra	love	dollar	buy
0	0.03	0.36	0.02	0.02
1	0.32	0.05	0.83	0.74

[0,1,0,1]

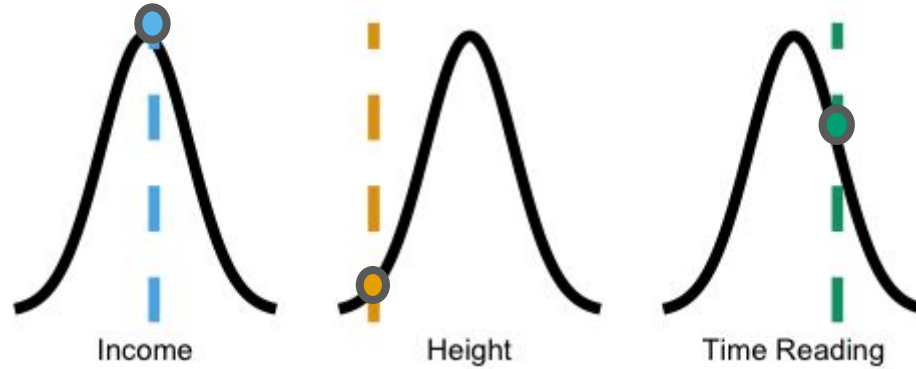
Bernoulli Naive Bayes



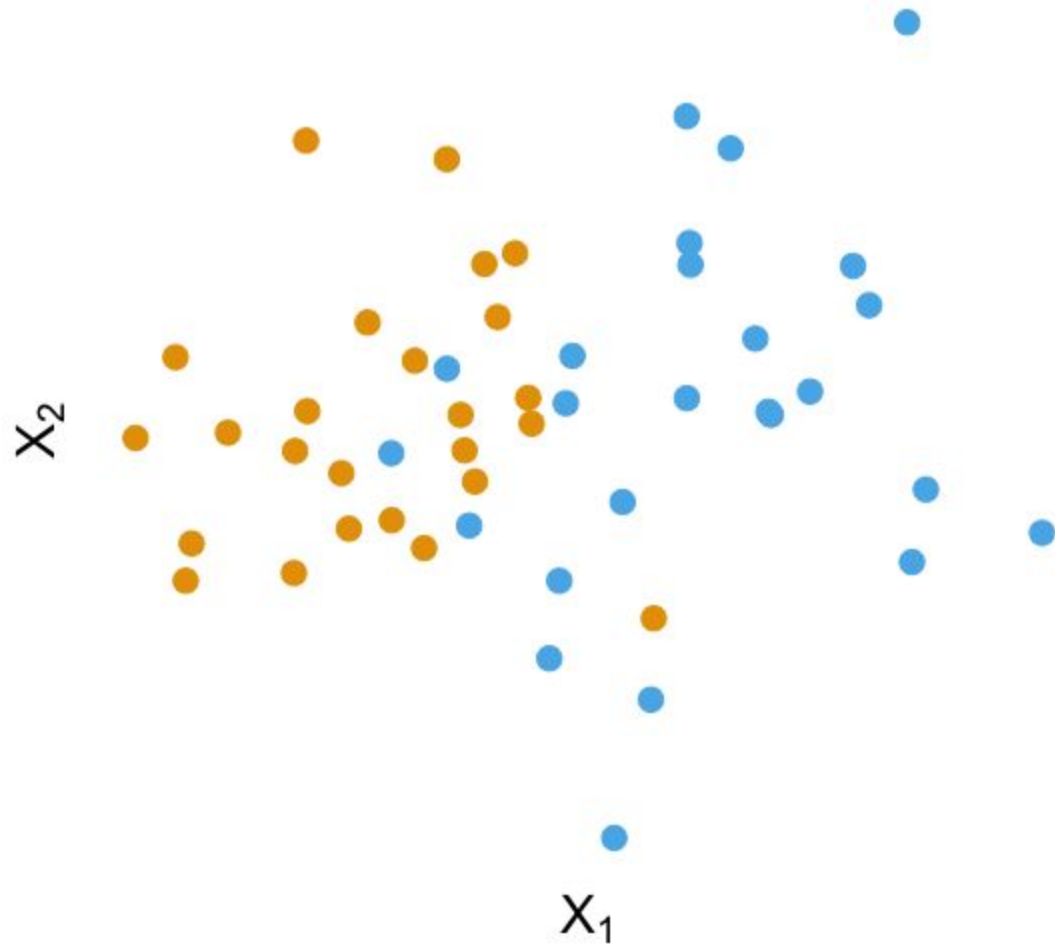
Categorical Naive Bayes

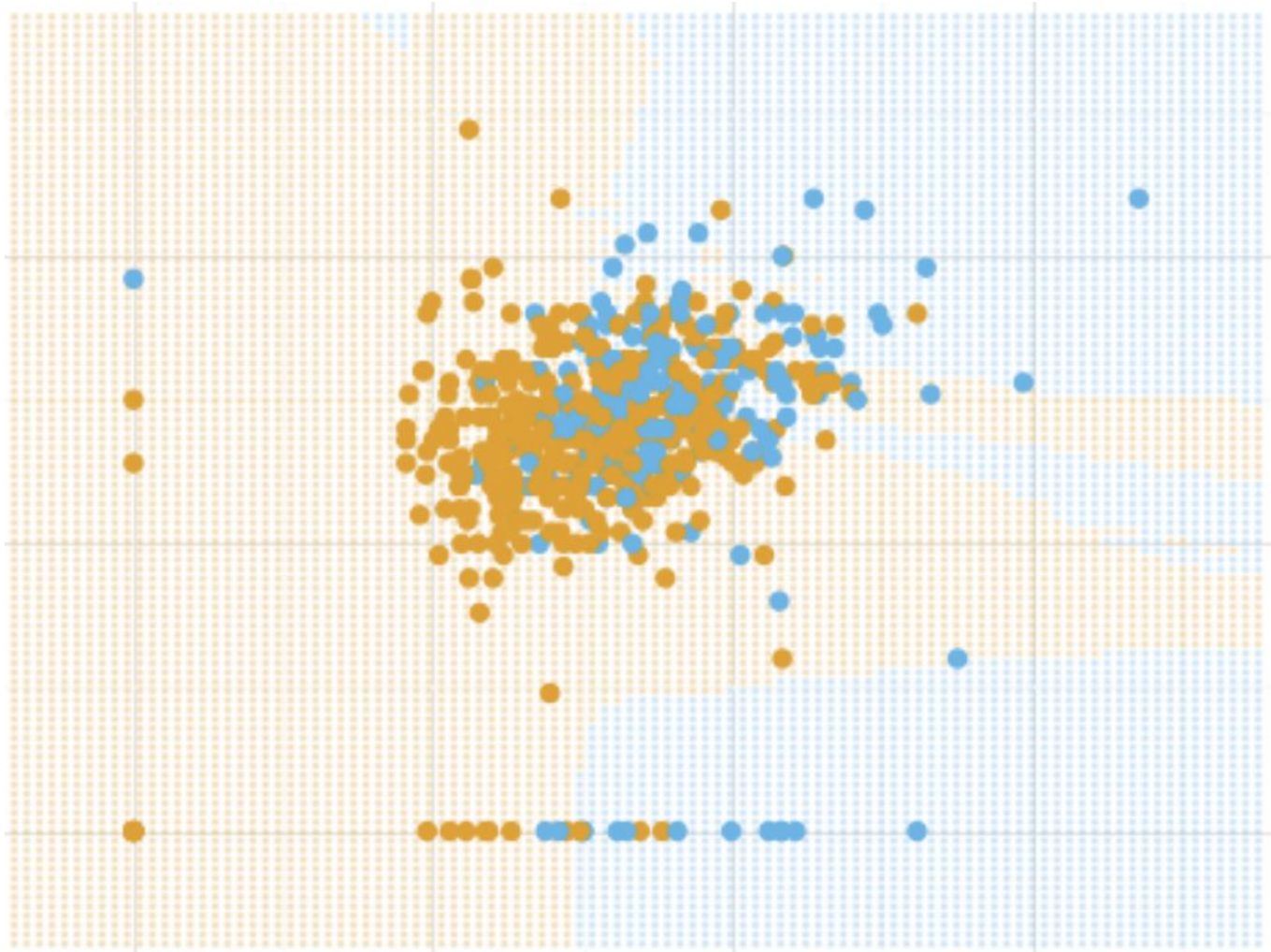


Gaussian Naive Bayes



K-Nearest Neighbors (KNN)





Hyperparameter Tuning

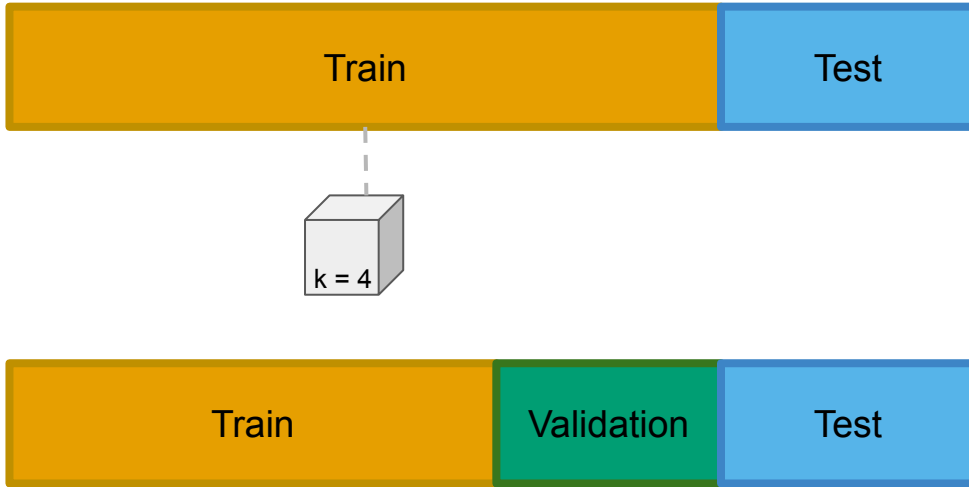
Train-Validation-Test

What k should I use 🙋

- Choose yourself
- Let the data decide

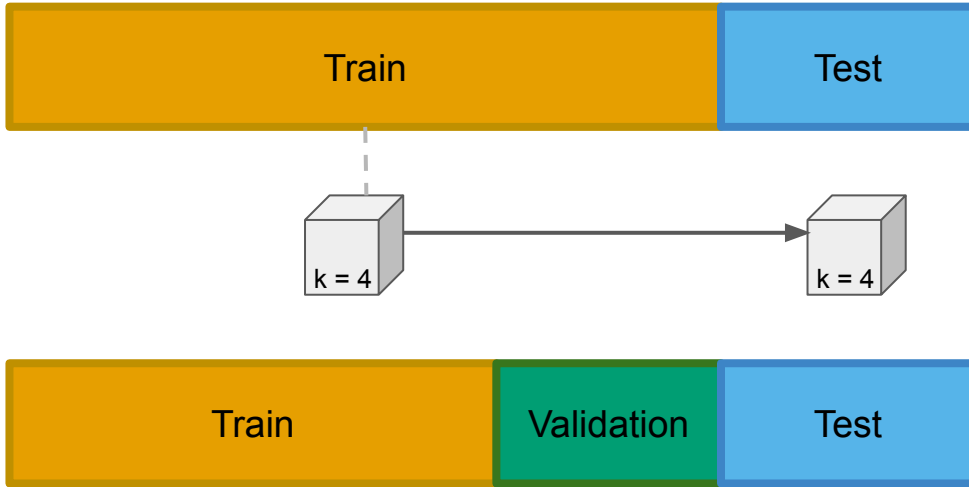
Train-Validation-Test

What k should I use 🙋



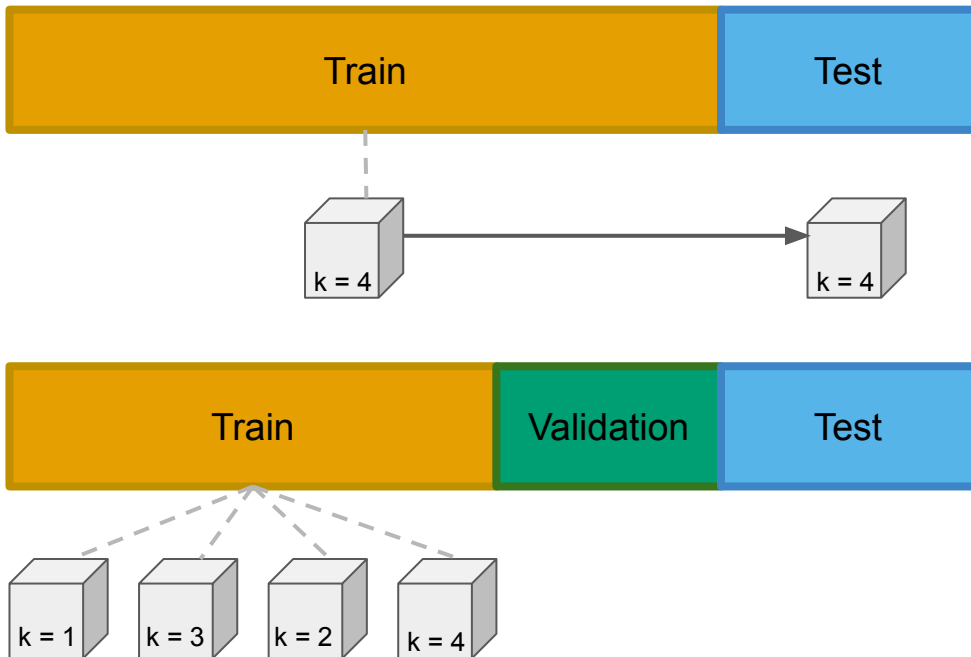
Train-Validation-Test

What k should I use 🙋



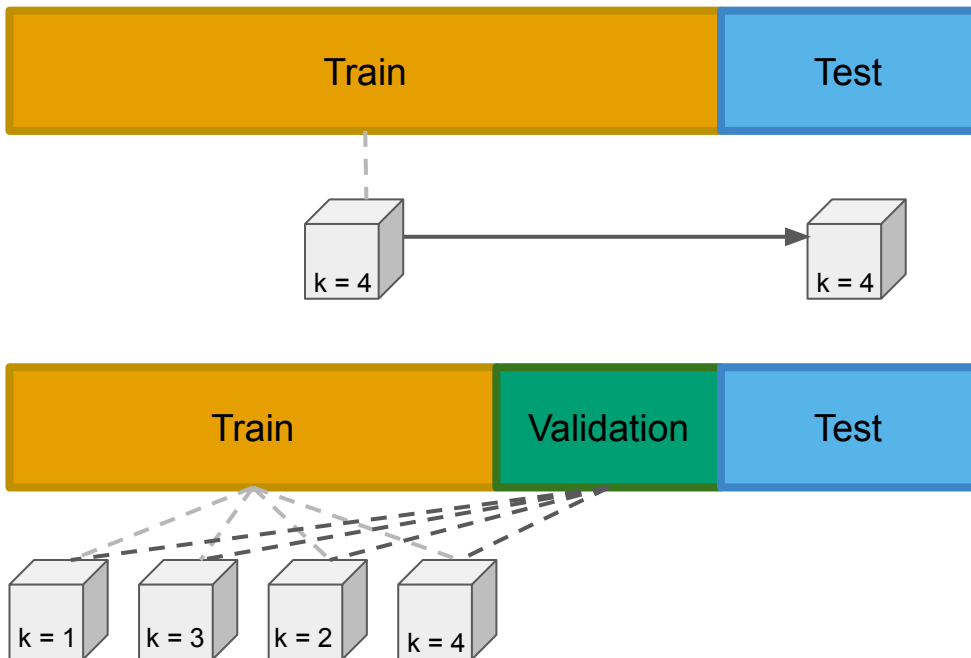
Train-Validation-Test

What k should I use 🙋



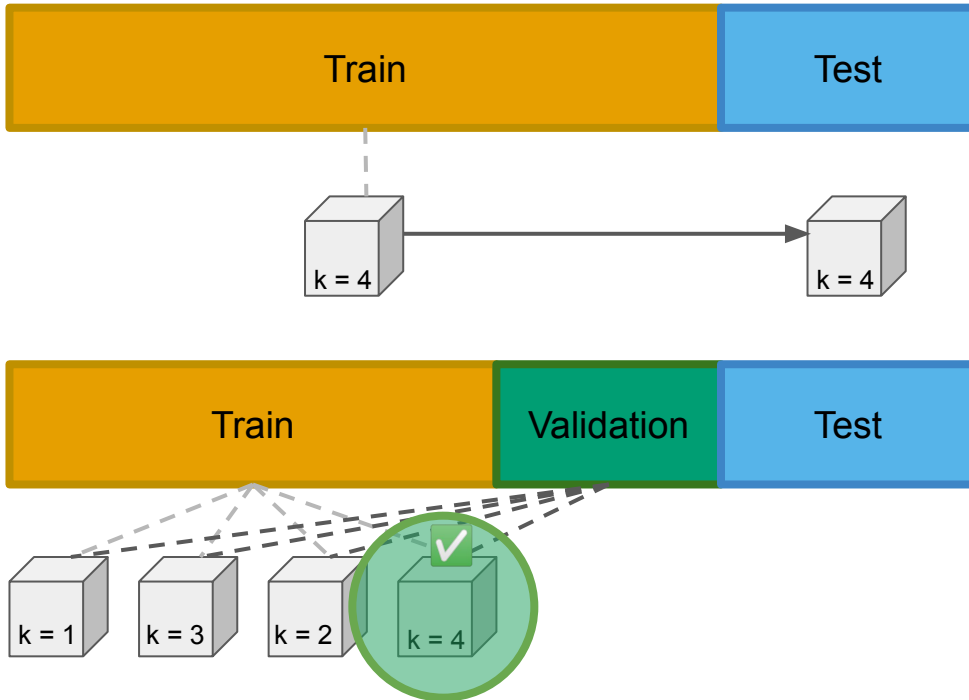
Train-Validation-Test

What k should I use 🙋



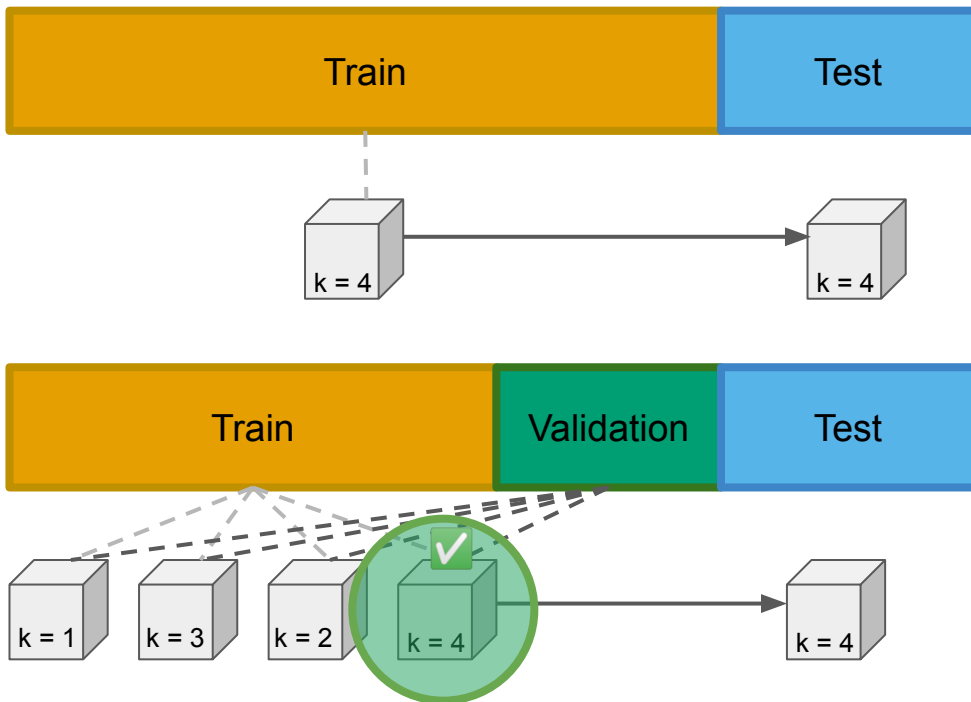
Train-Validation-Test

What k should I use 🙋

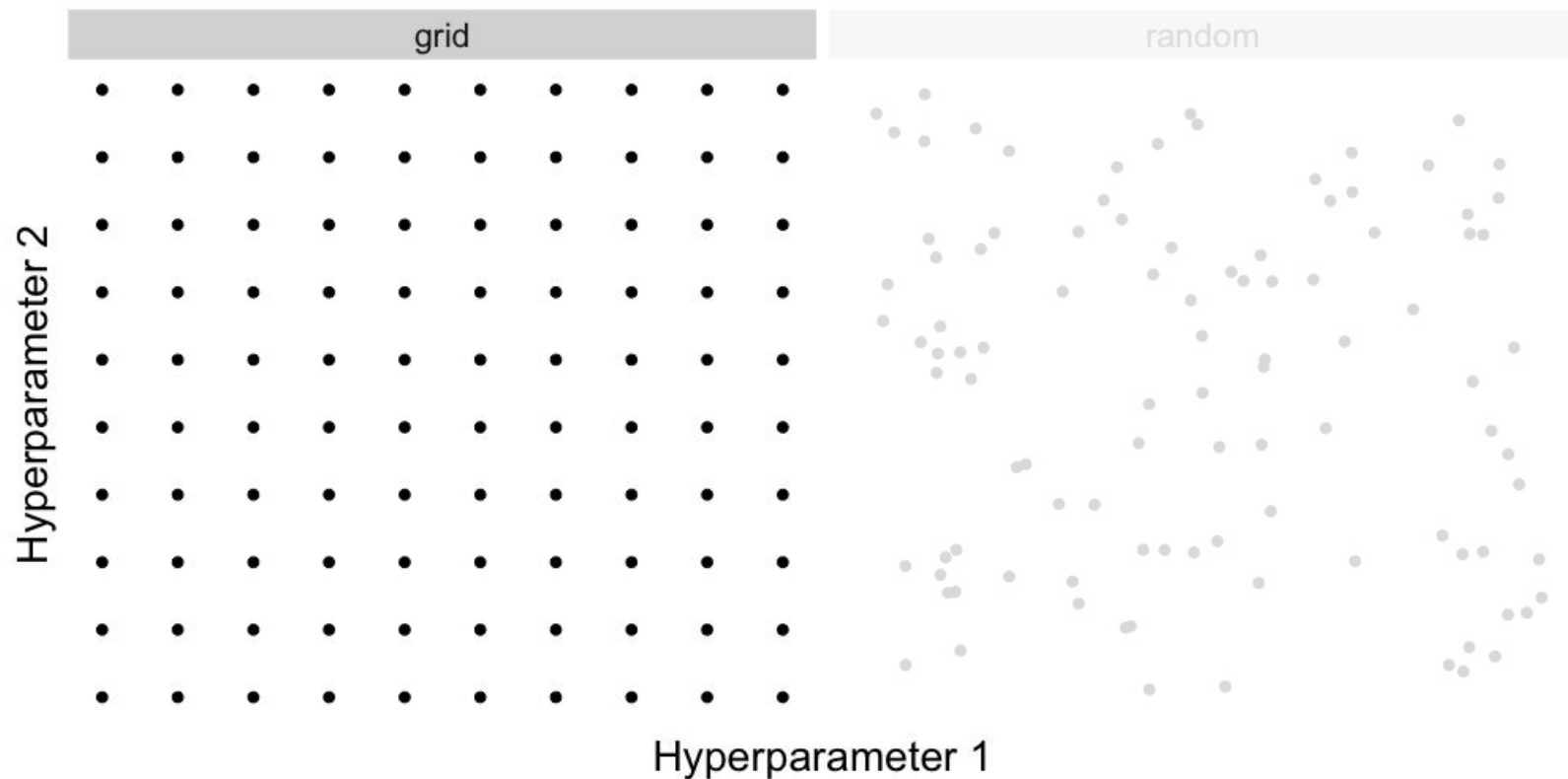


Train-Validation-Test

What k should I use 🙋



Grid Search



Random Search

