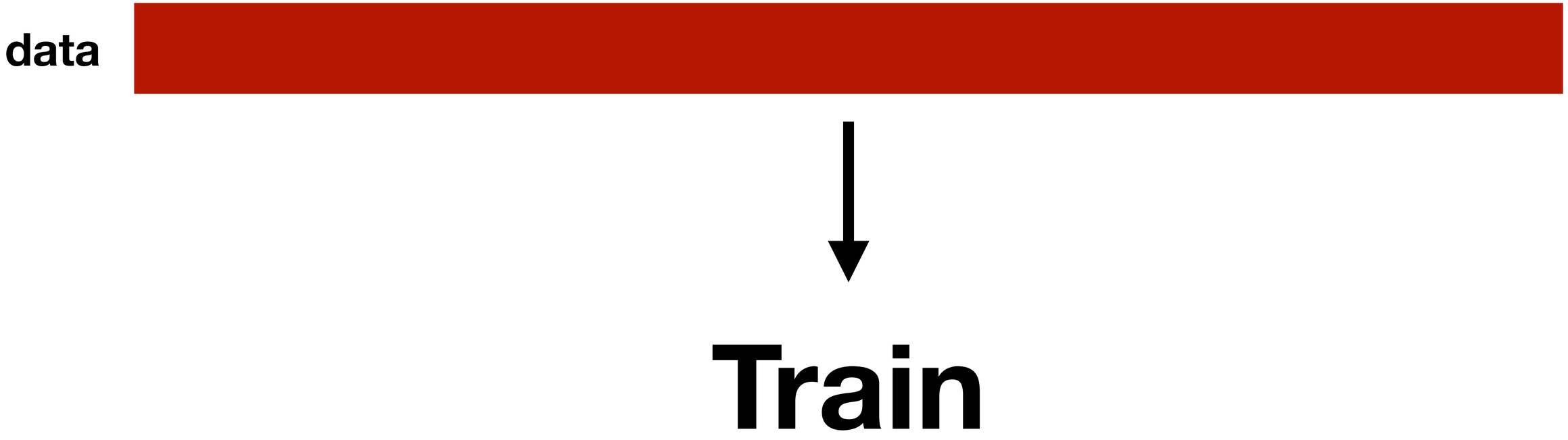
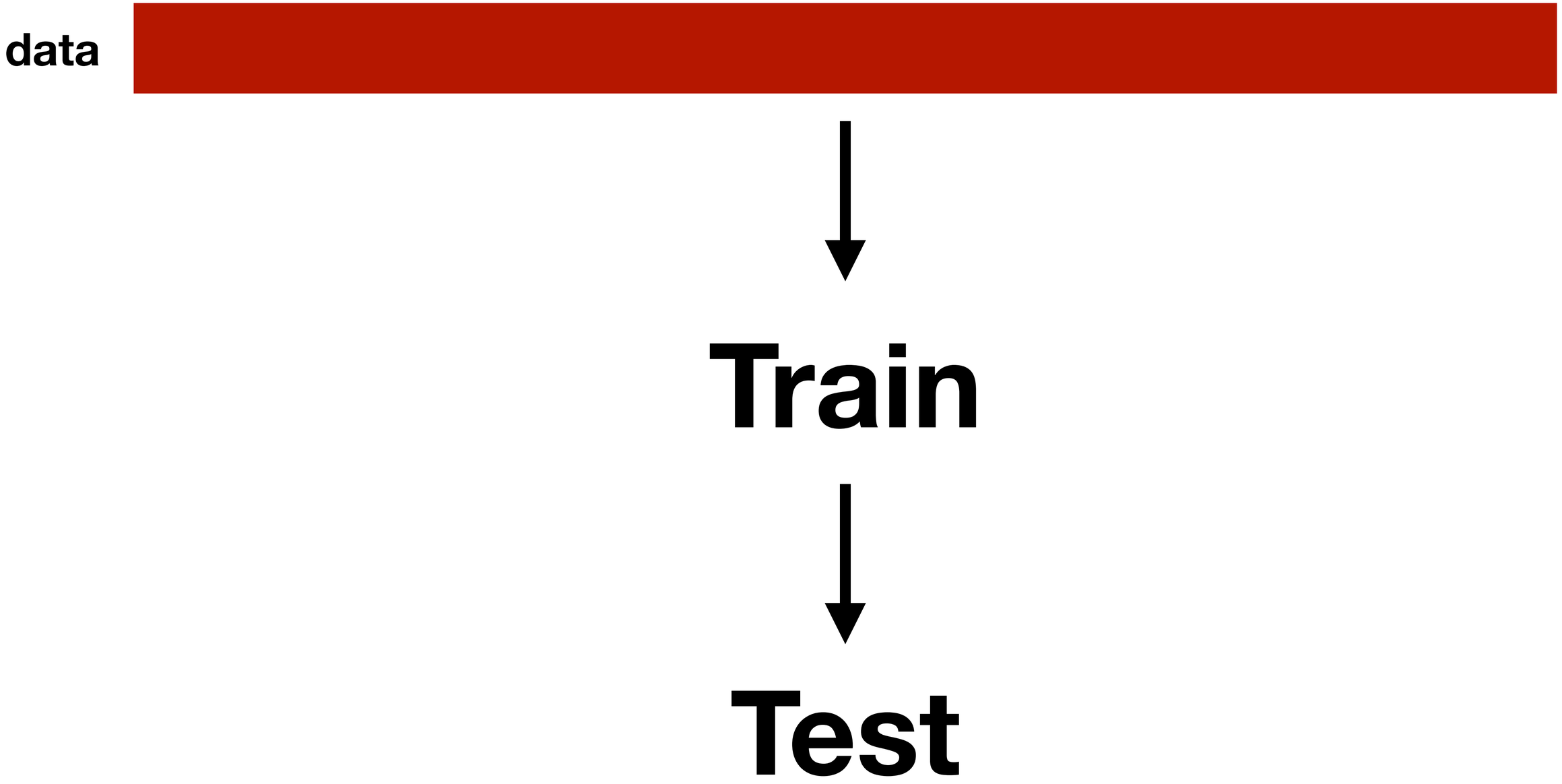
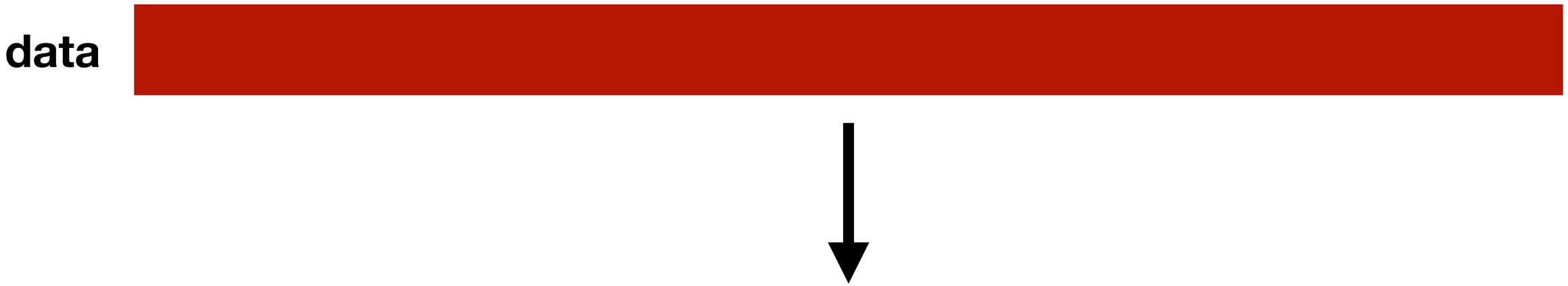
SC201

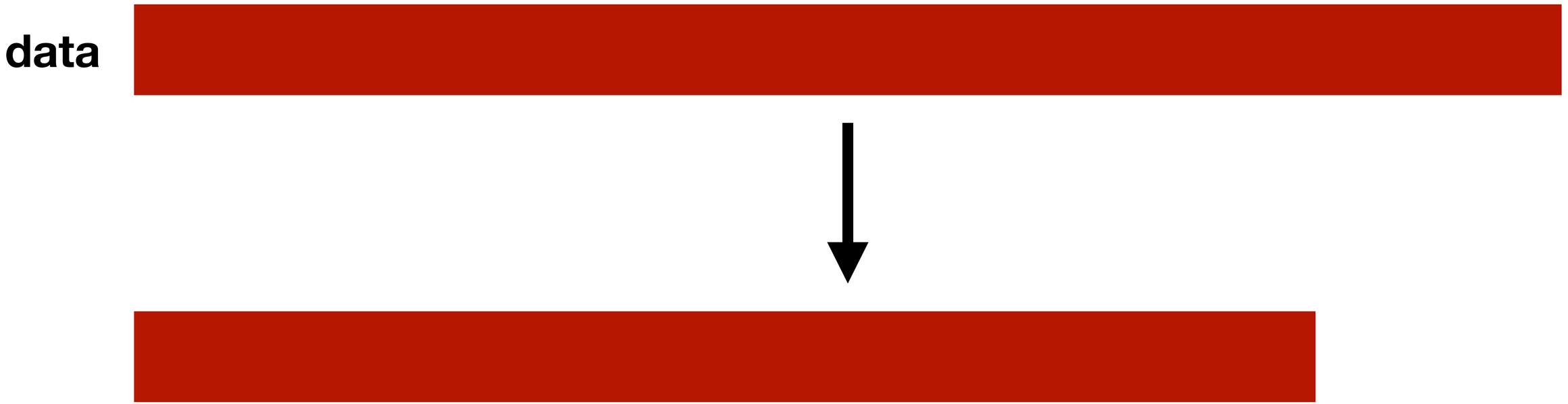
Lecture 5

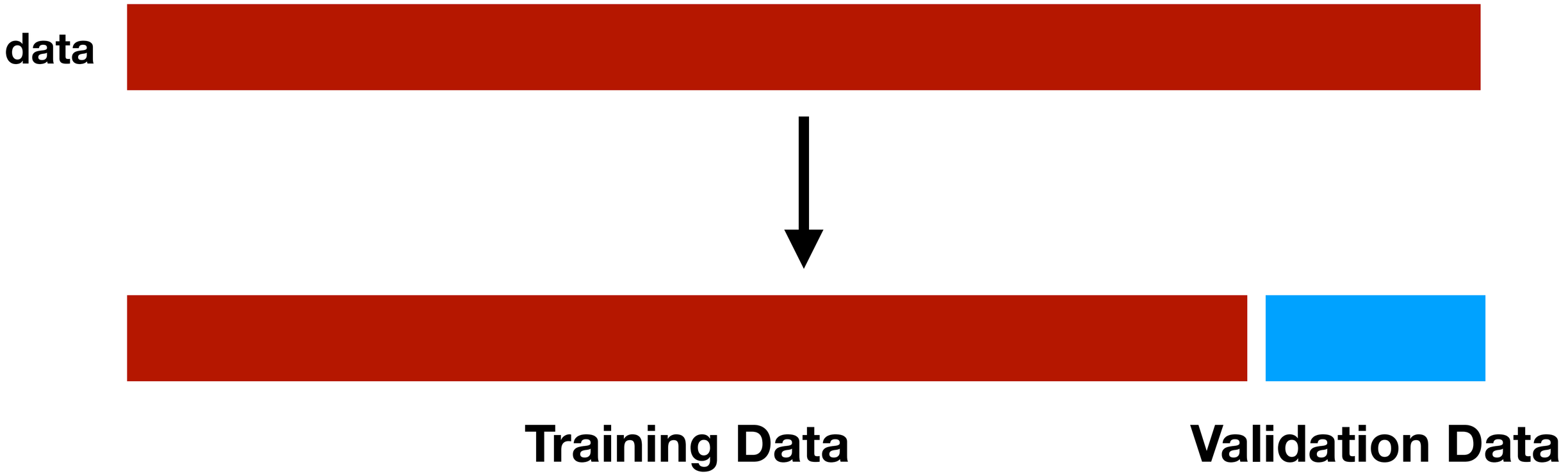
data

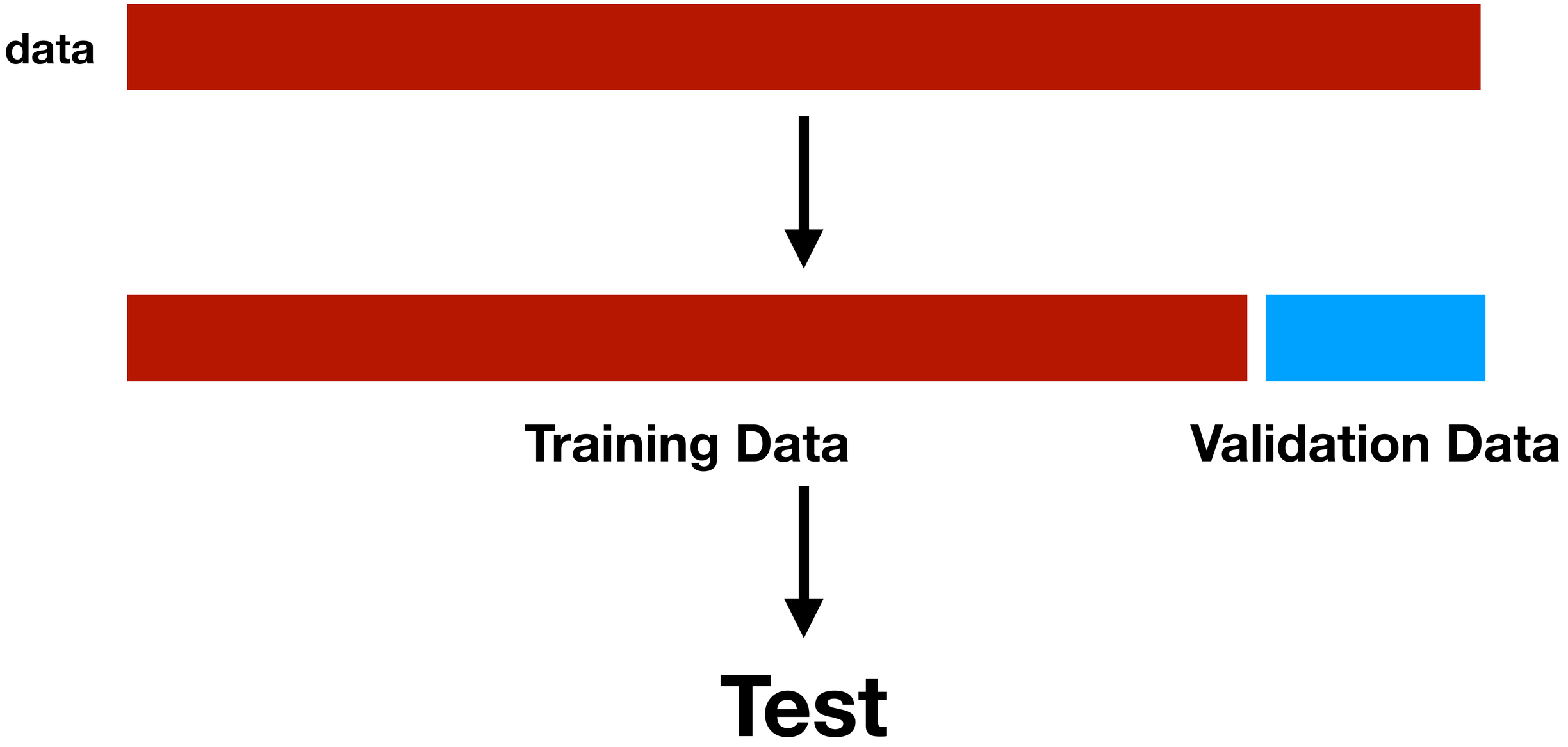

data
Train


data
Train
Test


data

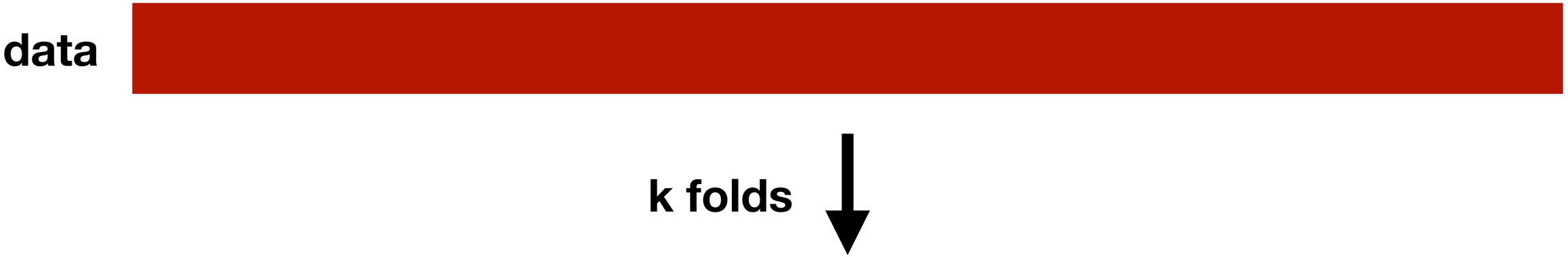

data
Training Data

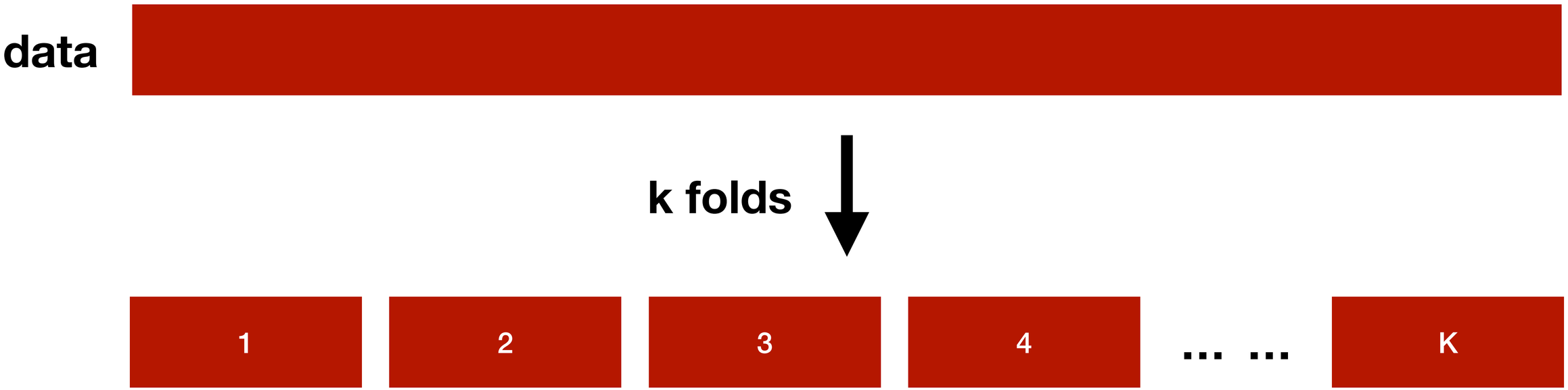

data
Training Data
Validation Data

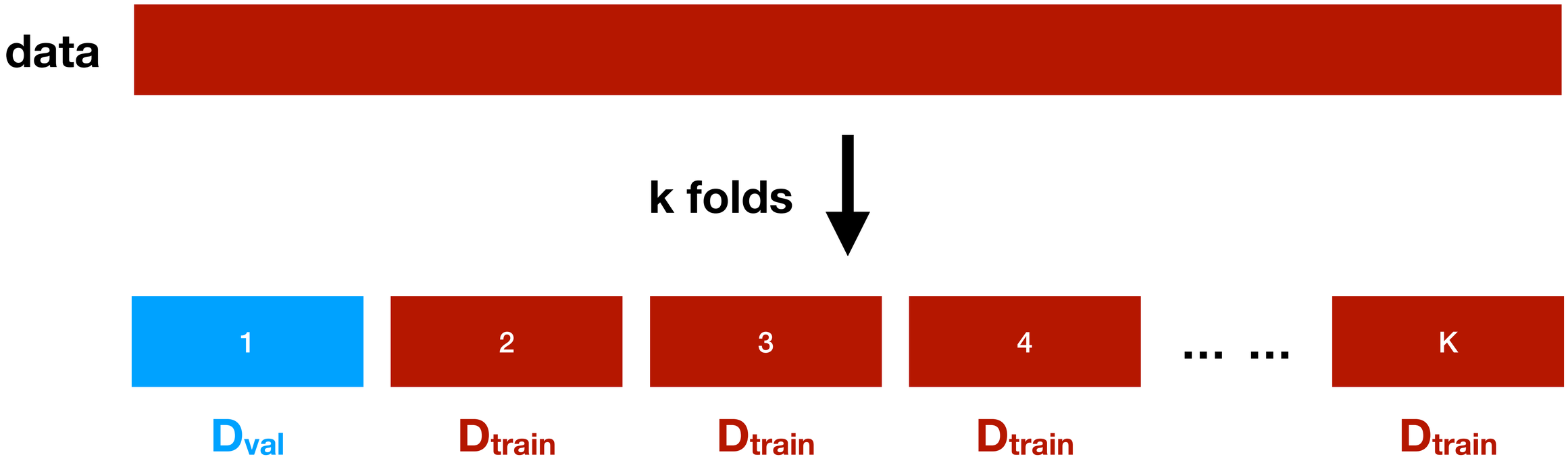

data
Training Data
Validation Data
Test

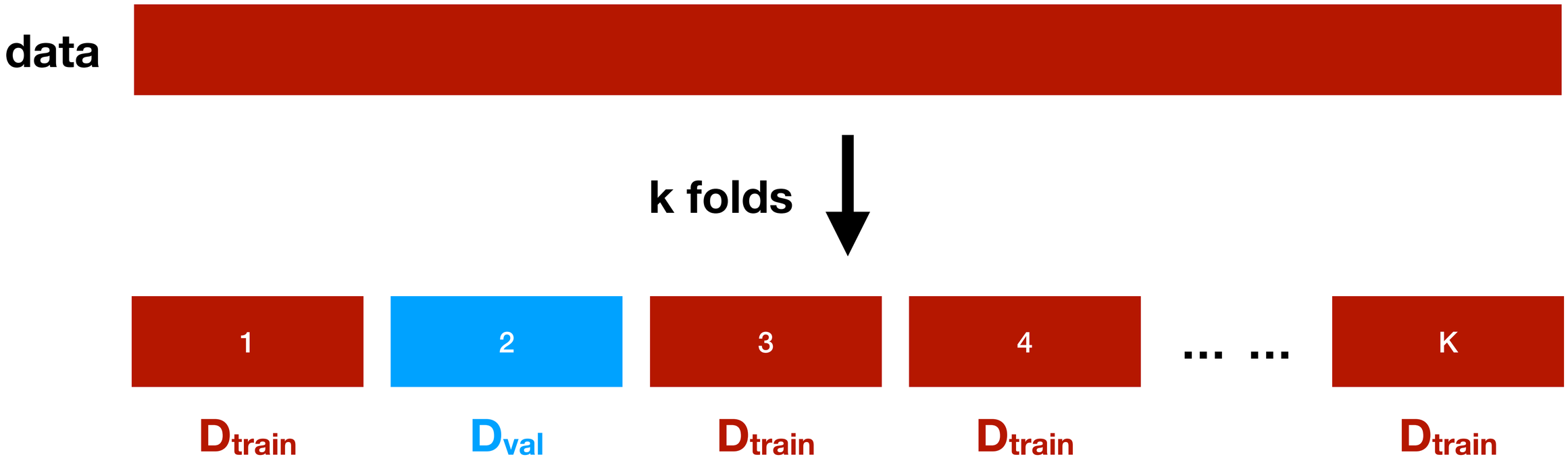
Cross Validation

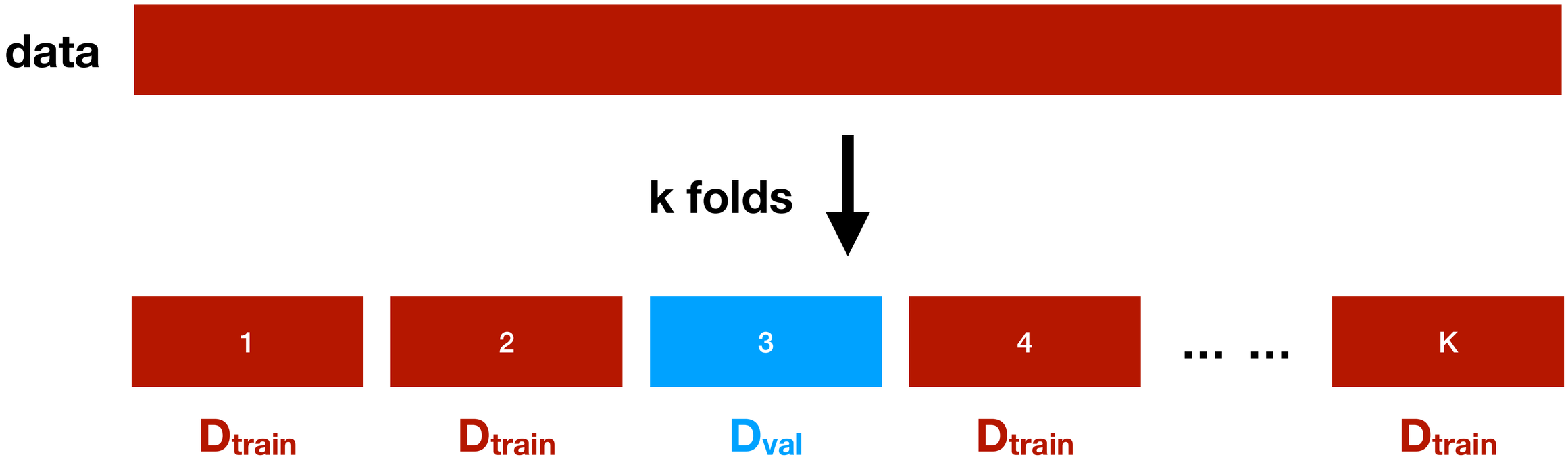
data

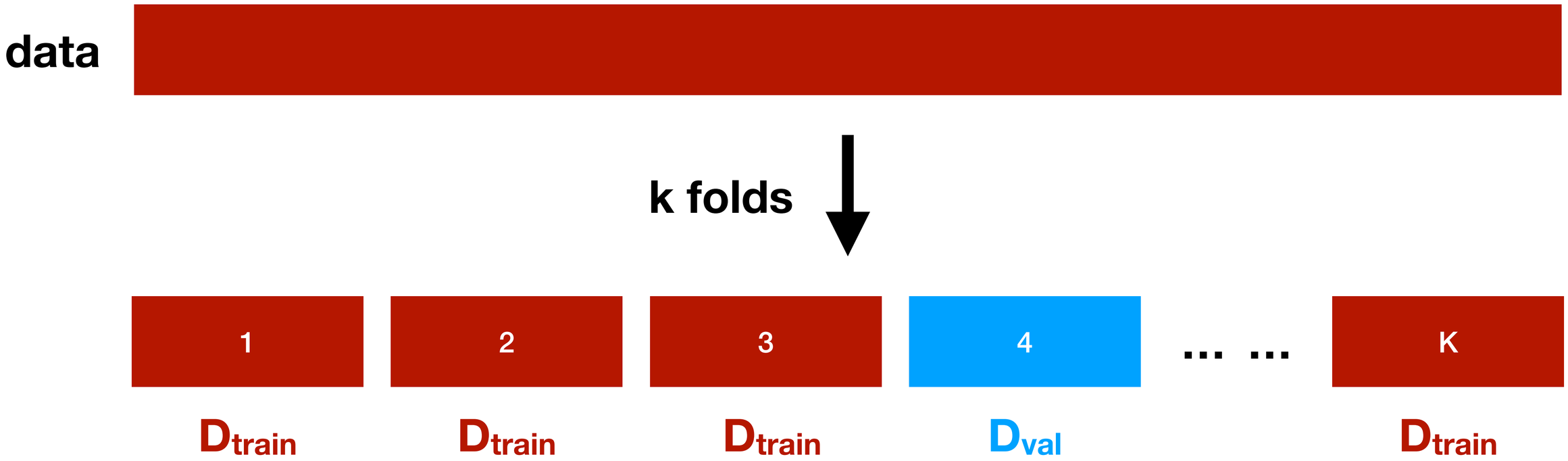

data
k folds

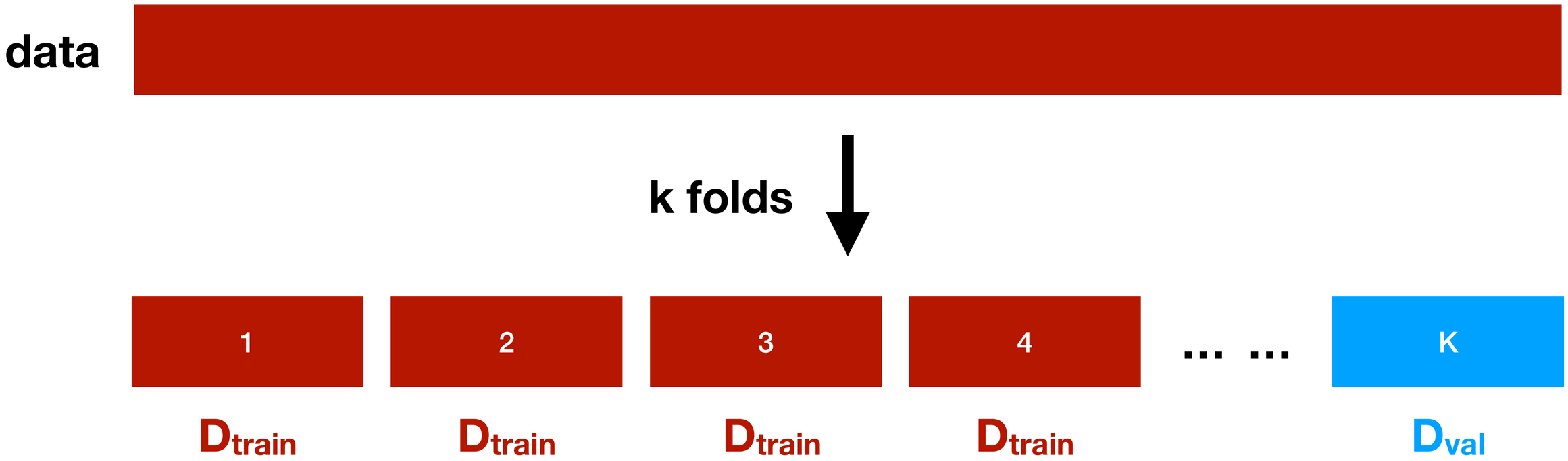

data
k folds
1
2
3
4
... ...
K

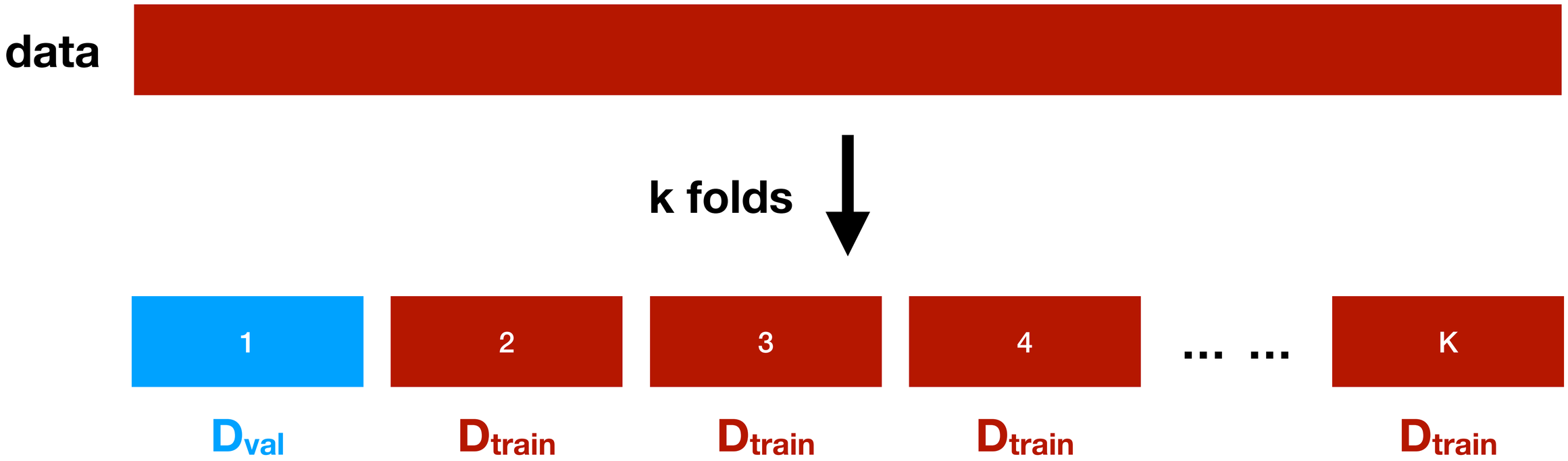

data
k folds
1
2
3
4
... ...
K
Dval
Dtrain
Dtrain
Dtrain
Dtrain

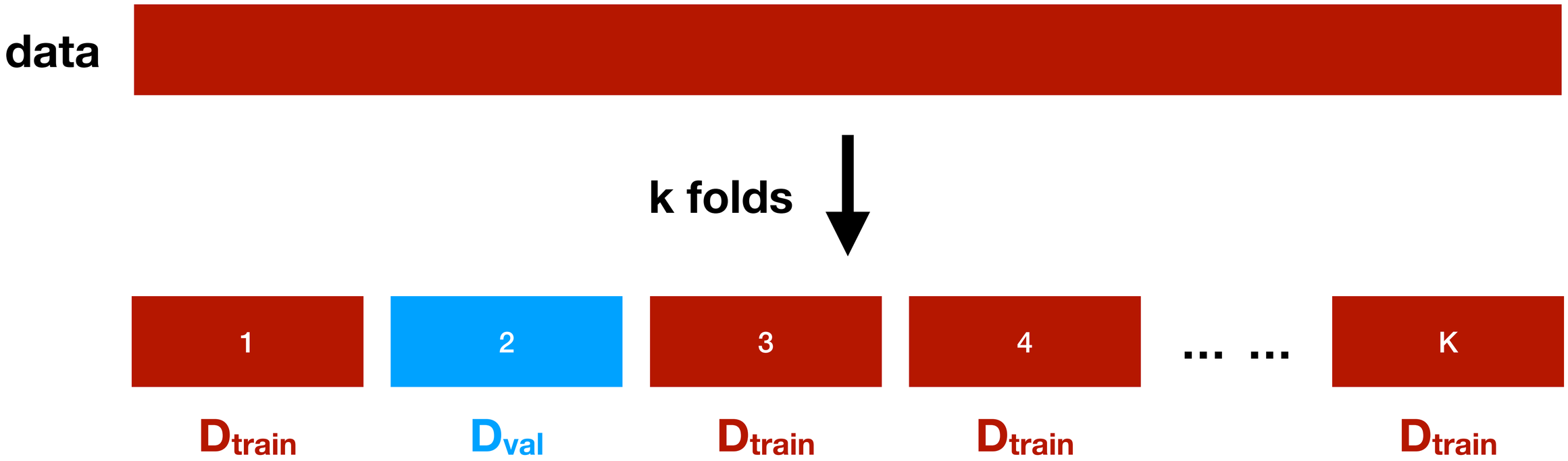

data
k folds
1
2
3
4
... ...
K
Dtrain
Dval
Dtrain
Dtrain
Dtrain

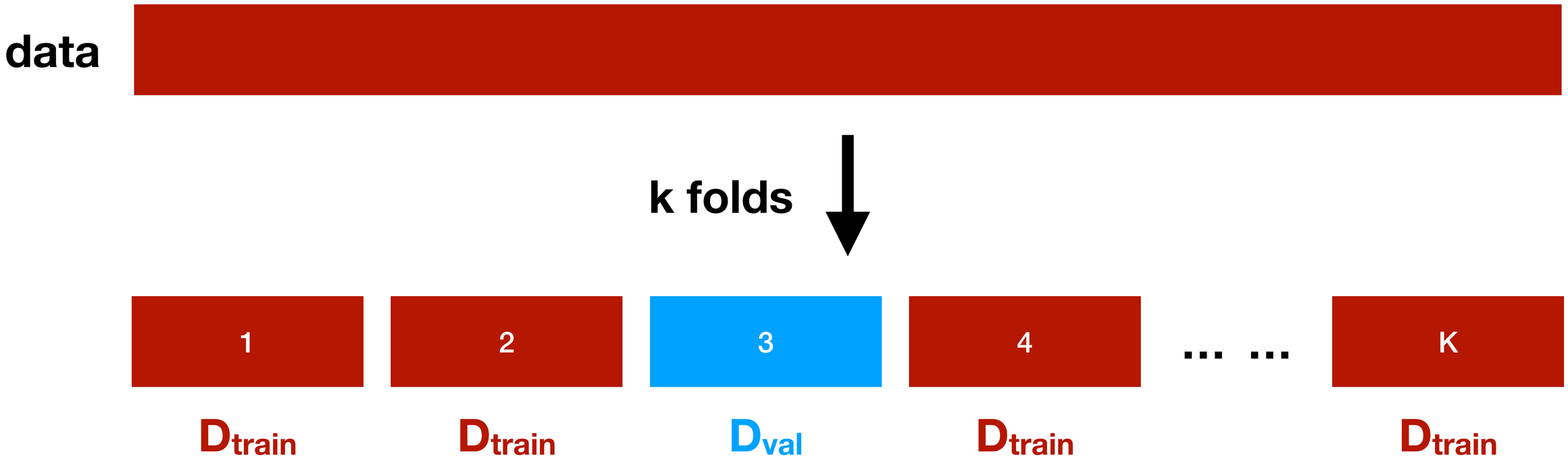

data
k folds
1
2
3
4
... ...
K
Dtrain
Dtrain
Dval
Dtrain
Dtrain

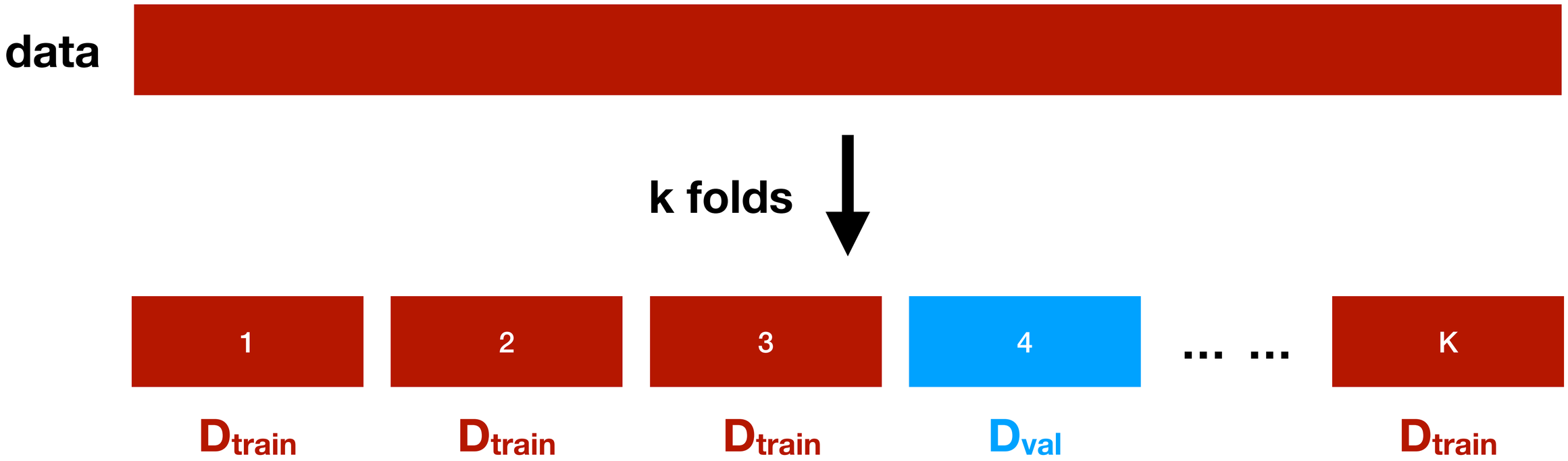

data
k folds
1
2
3
4
... ...
K
Dtrain
Dtrain
Dtrain
Dval
Dtrain

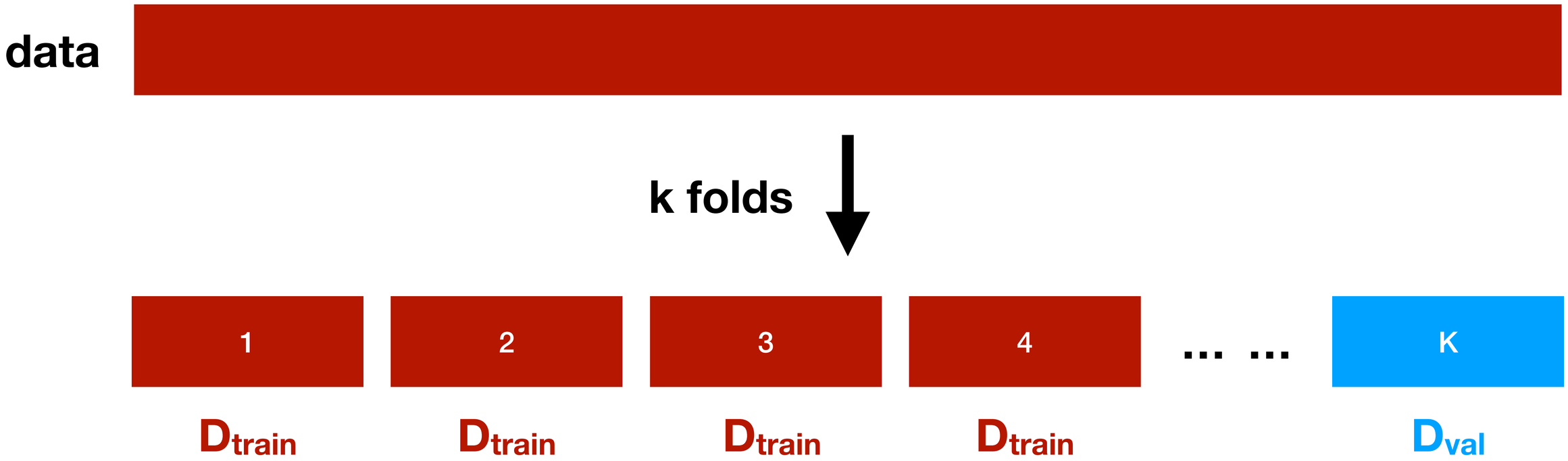

data
k folds
1
2
3
4
... ...
K
Dtrain
Dtrain
Dtrain
Dtrain
Dval


data
k folds
1
2
3
4
... ...
K
Dval
Dtrain
Dtrain
Dtrain
Dtrain


data
k folds
1
2
3
4
... ...
K
Dtrain
Dval
Dtrain
Dtrain
Dtrain


data
k folds
1
2
3
4
... ...
K
Dtrain
Dtrain
Dval
Dtrain
Dtrain


data
k folds
1
2
3
4
... ...
K
Dtrain
Dtrain
Dtrain
Dval
Dtrain


data
k folds
1
2
3
4
... ...
K
Dtrain
Dtrain
Dtrain
Dtrain
Dval

房地產公司要您建造一個「房價預測模型」！

如果您的模型估價比市價高，沒有顧客會想要跟你們家買房

但如果估價比市價低，公司利潤又會大受影響

因此最好的辦法就是讓模型預測出來的結果「比正確數值高一點」

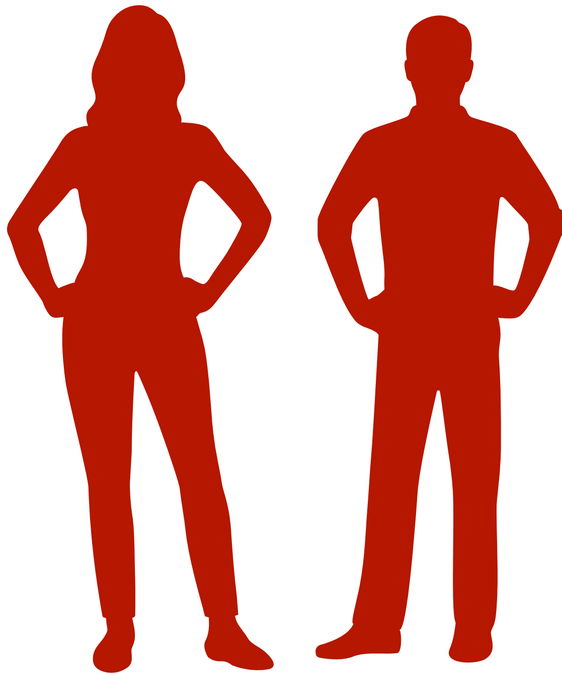
這樣還可以有降價的空間（如果比市價低，買走了，就沒有了...）

假設 y_i 是 True Label, h_i 是模型預測的結果

請問您要怎麼調整 Loss Function (L)

(對 $L = (h_i - y_i)^2$ 做調整)

讓模型 train 完之後產出的結果就會比正確數值高？



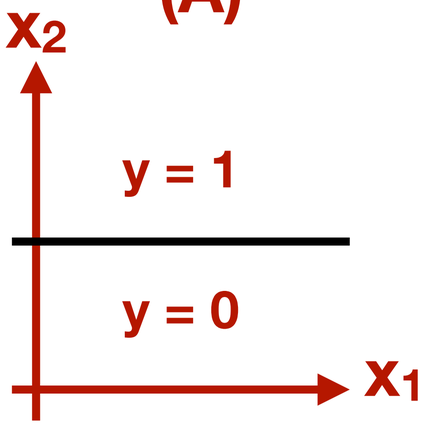
Suppose you train a logistic regression classifier and your hypothesis function h is

$$h = \sigma(\theta_0 + \theta_1 x_1 + \theta_2 x_2)$$

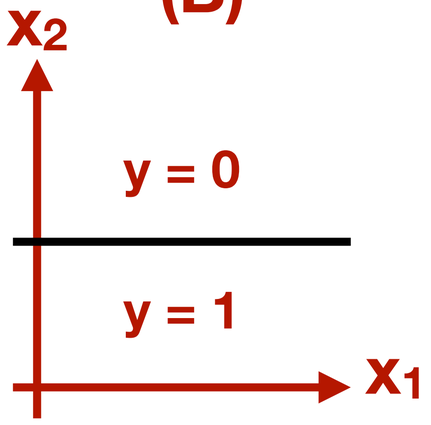
where $\theta_0 = 6$; $\theta_1 = 0$; $\theta_2 = -1$

Which of the following figure will represent the decision boundary as given by above classifier?

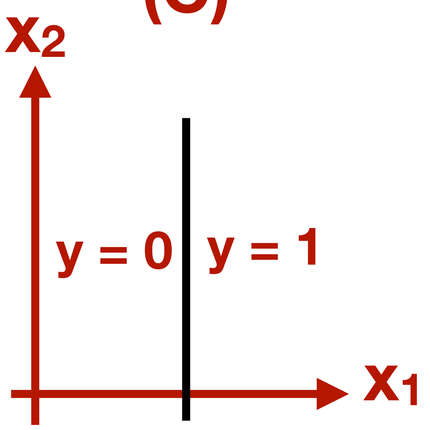
(A)



(B)



(C)



(D)