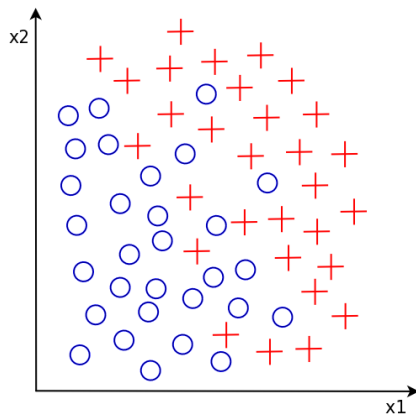


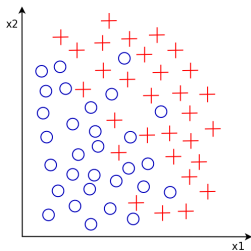
Cross-validation

Classification task

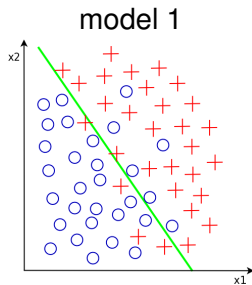
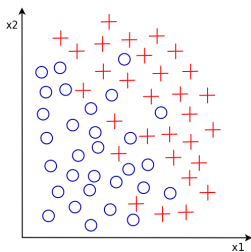


- task: classify future independent observations drawn from the same distribution
- how to choose a (classification) model?

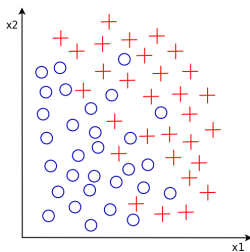
Model selection (1)



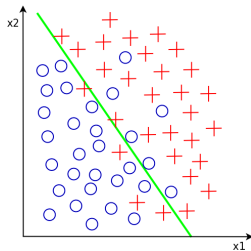
Model selection (1)



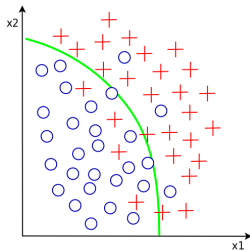
Model selection (1)



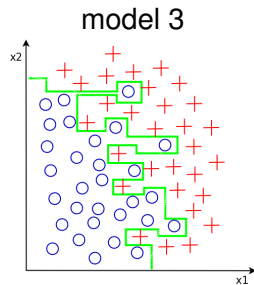
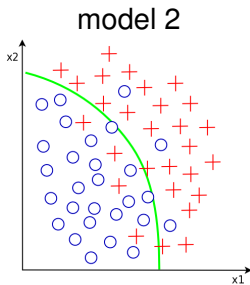
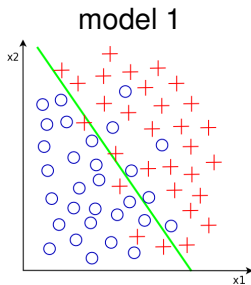
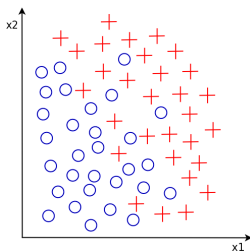
model 1



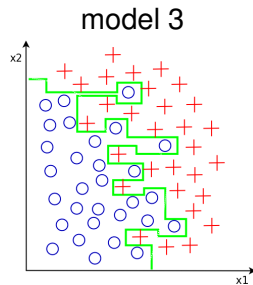
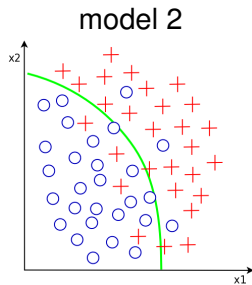
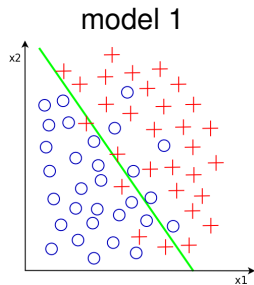
model 2



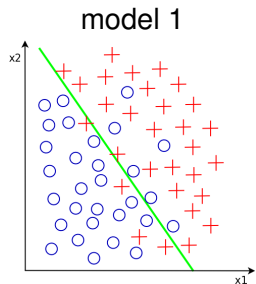
Model selection (1)



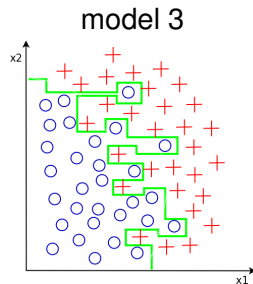
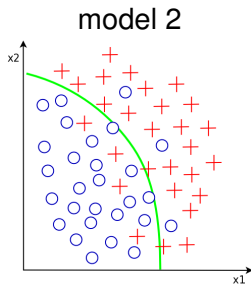
Model selection (2)



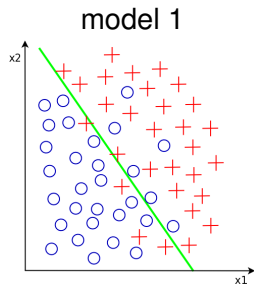
Model selection (2)



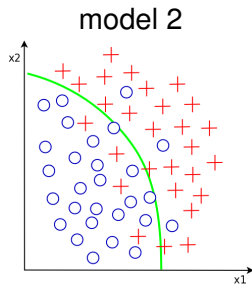
10 misclassifications



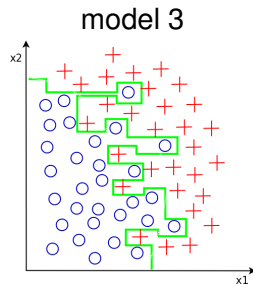
Model selection (2)



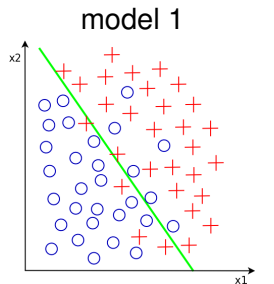
10 misclassifications



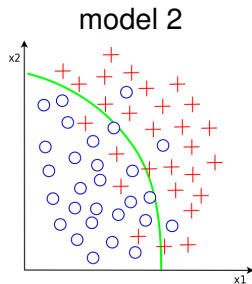
7 missclass.



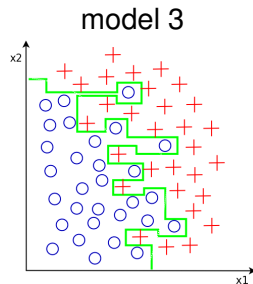
Model selection (2)



10 misclassifications

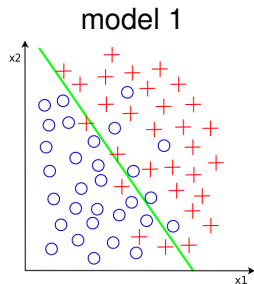


7 missclass.

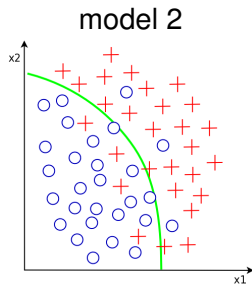


0 missclass.

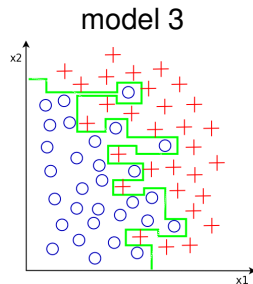
Model selection (2)



10 misclassifications
2 parameters

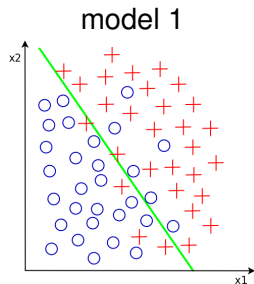


7 missclass.

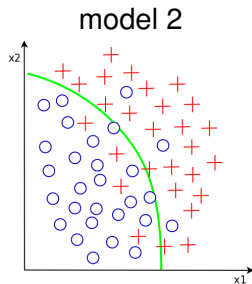


0 missclass.

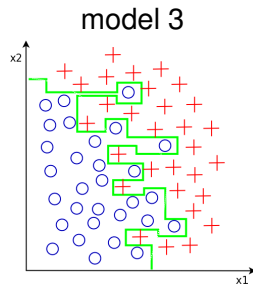
Model selection (2)



10 misclassifications
2 parameters

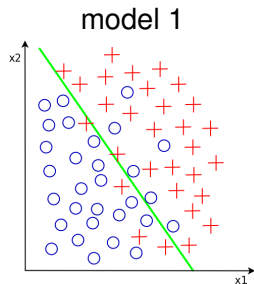


7 misclass.
4 parameters

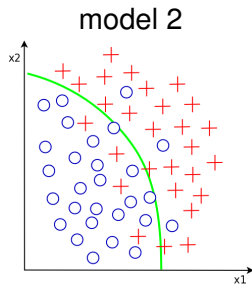


0 misclass.

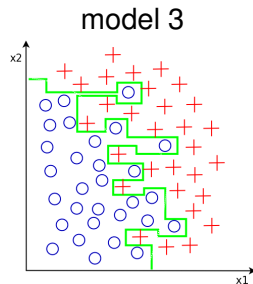
Model selection (2)



10 misclassifications
2 parameters

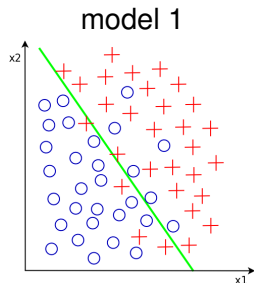


7 missclass.
4 parameters

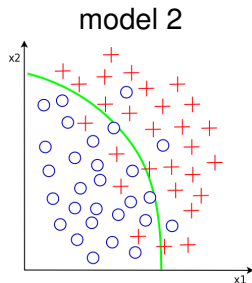


0 missclass.
72 parameters

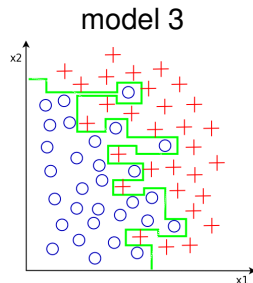
Model selection (2)



10 missclassifications
2 parameters



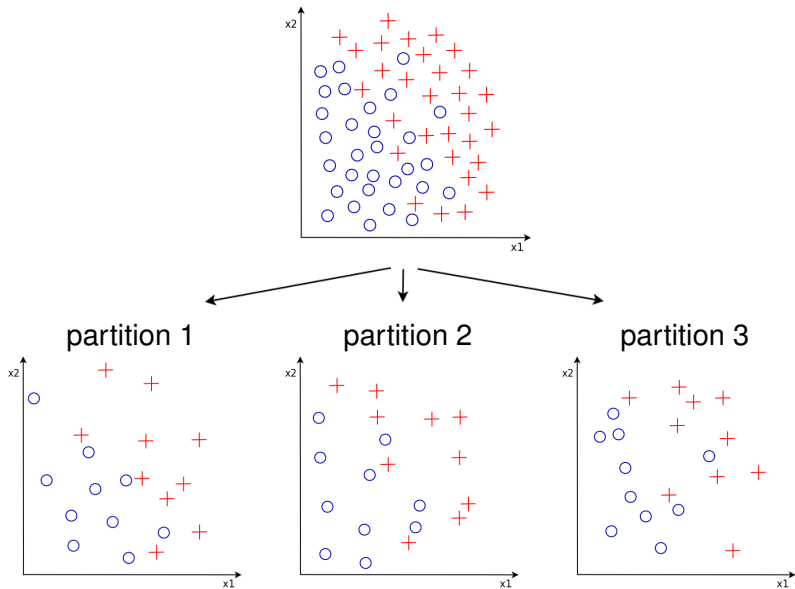
7 missclass.
4 parameters



0 missclass.
72 parameters

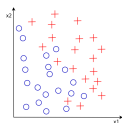
- which model is the best?

Cross-validation: data splitting



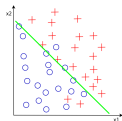
Cross-validating model 1

training set	train error	testing set	test error
<hr/>			
partition 1+2			



Cross-validating model 1

training set	train error	testing set	test error
partition 1+2			



4/20

Cross-validating model 1

training set

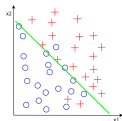
train error

testing set

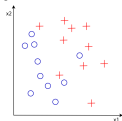
test error

partition 1+2

partition 3

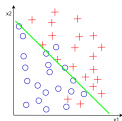


4/20



Cross-validating model 1

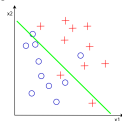
training set
partition 1+2



4/20

train error

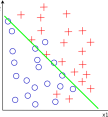
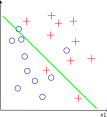
testing set
partition 3

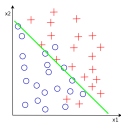


4/10

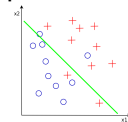
test error

Cross-validating model 1

training set	train error	testing set	test error
partition 1+2		partition 3	
	4/20		4/10
partition 1+3			

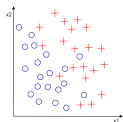


4/20

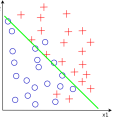
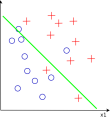
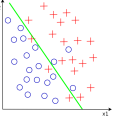


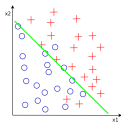
4/10

partition 1+3

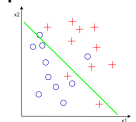


Cross-validating model 1

training set	train error	testing set	test error
partition 1+2		partition 3	
	4/20		4/10
partition 1+3			
	6/20		

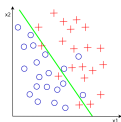


4/20



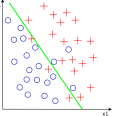
4/10

partition 1+3

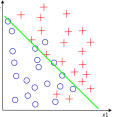
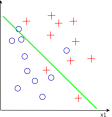
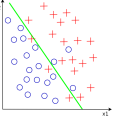
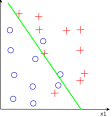


6/20

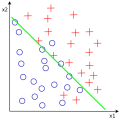
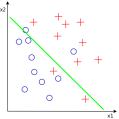
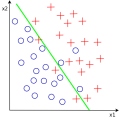
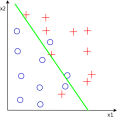
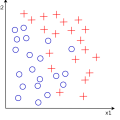
Cross-validating model 1

training set	train error	testing set	test error
partition 1+2 	4/20	partition 3 	4/10
partition 1+3 	6/20	partition 2 	

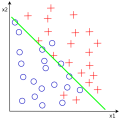
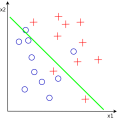
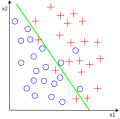
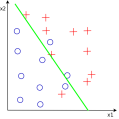
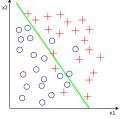
Cross-validating model 1

training set	train error	testing set	test error
partition 1+2 	4/20	partition 3 	4/10
partition 1+3 	6/20	partition 2 	3/10

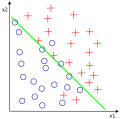
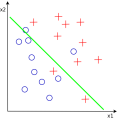
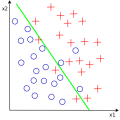
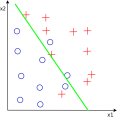
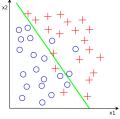
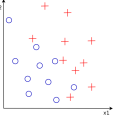
Cross-validating model 1

training set	train error	testing set	test error
partition 1+2 	4/20	partition 3 	4/10
partition 1+3 	6/20	partition 2 	3/10
partition 2+3 			

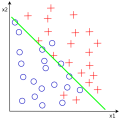
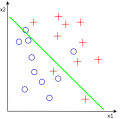
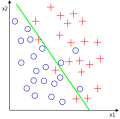
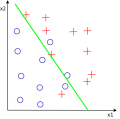
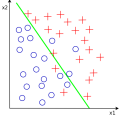
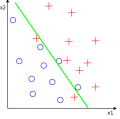
Cross-validating model 1

training set	train error	testing set	test error
partition 1+2 	4/20	partition 3 	4/10
partition 1+3 	6/20	partition 2 	3/10
partition 2+3 	5/20		

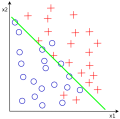
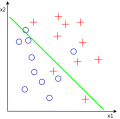
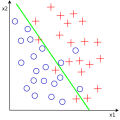
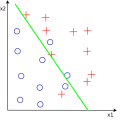
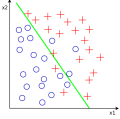
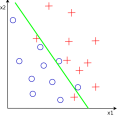
Cross-validating model 1

training set	train error	testing set	test error
partition 1+2 	4/20	partition 3 	4/10
partition 1+3 	6/20	partition 2 	3/10
partition 2+3 	5/20	partition 1 	

Cross-validating model 1

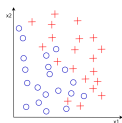
training set	train error	testing set	test error
partition 1+2 	4/20	partition 3 	4/10
partition 1+3 	6/20	partition 2 	3/10
partition 2+3 	5/20	partition 1 	4/10

Cross-validating model 1

training set	train error	testing set	test error	mean test error
partition 1+2 	4/20	partition 3 	4/10	
partition 1+3 	6/20	partition 2 	3/10	11/30
partition 2+3 	5/20	partition 1 	4/10	

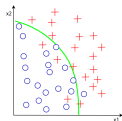
Cross-validating model 2

training set	train error	testing set	test error
<hr/>			
partition 1+2			



Cross-validating model 2

training set	train error	testing set	test error
partition 1+2			



4/20

Cross-validating model 2

training set

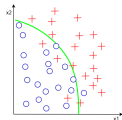
train error

testing set

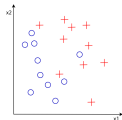
test error

partition 1+2

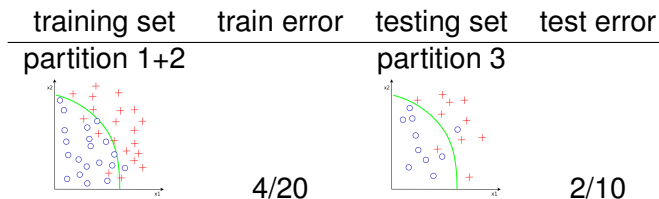
partition 3



4/20

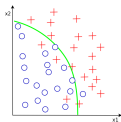


Cross-validating model 2



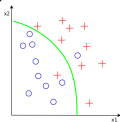
Cross-validating model 2

training set
partition 1+2



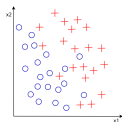
4/20

testing set
partition 3

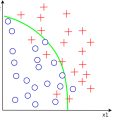
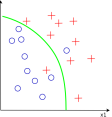
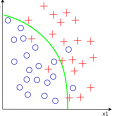


2/10

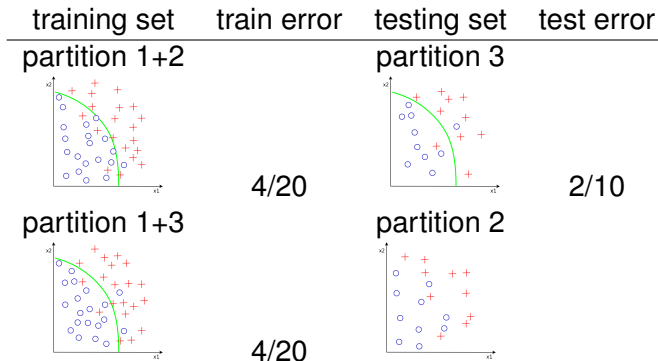
partition 1+3



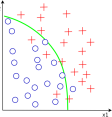
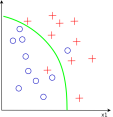
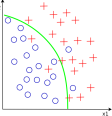
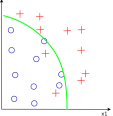
Cross-validating model 2

training set	train error	testing set	test error
partition 1+2		partition 3	
	4/20		2/10
partition 1+3			
	4/20		

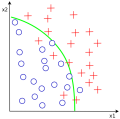
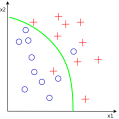
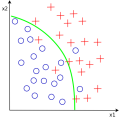
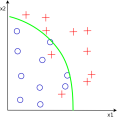
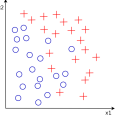
Cross-validating model 2



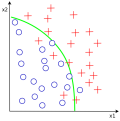
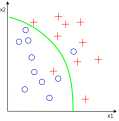
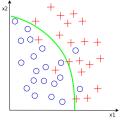
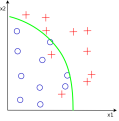
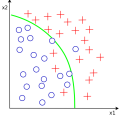
Cross-validating model 2

training set	train error	testing set	test error
partition 1+2 	4/20	partition 3 	2/10
partition 1+3 	4/20	partition 2 	2/10

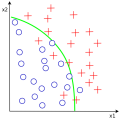
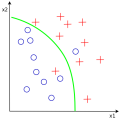
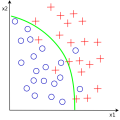
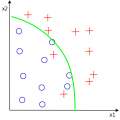
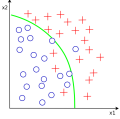
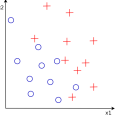
Cross-validating model 2

training set	train error	testing set	test error
partition 1+2 	4/20	partition 3 	2/10
partition 1+3 	4/20	partition 2 	2/10
partition 2+3 			

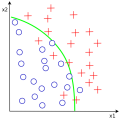
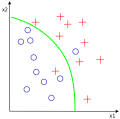
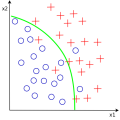
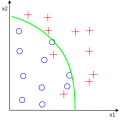
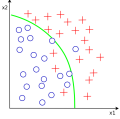
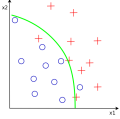
Cross-validating model 2

training set	train error	testing set	test error
partition 1+2 	4/20	partition 3 	2/10
partition 1+3 	4/20	partition 2 	2/10
partition 2+3 	4/20		

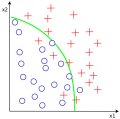
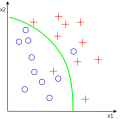
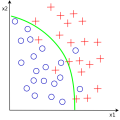
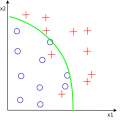
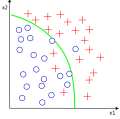
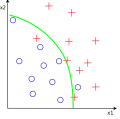
Cross-validating model 2

training set	train error	testing set	test error
partition 1+2 	4/20	partition 3 	2/10
partition 1+3 	4/20	partition 2 	2/10
partition 2+3 	4/20	partition 1 	

Cross-validating model 2

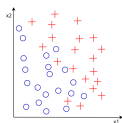
training set	train error	testing set	test error
partition 1+2		partition 3	
	4/20		2/10
partition 1+3		partition 2	
	4/20		2/10
partition 2+3		partition 1	
	4/20		2/10

Cross-validating model 2

training set	train error	testing set	test error	mean test error
partition 1+2		partition 3		
	4/20		2/10	
partition 1+3		partition 2		
	4/20		2/10	6/30
partition 2+3		partition 1		
	4/20		2/10	

Cross-validating model 3

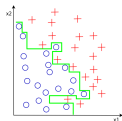
training set	train error	testing set	test error
<hr/>			
partition 1+2			



Cross-validating model 3

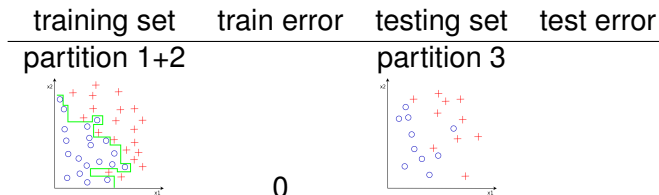
training set	train error	testing set	test error
--------------	-------------	-------------	------------

partition 1+2



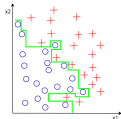
0

Cross-validating model 3

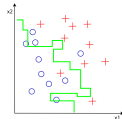


Cross-validating model 3

training set	train error	testing set	test error
partition 1+2	0	partition 3	5/10

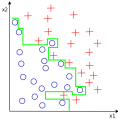
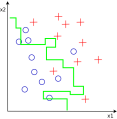


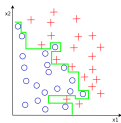
0



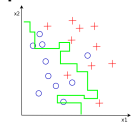
5/10

Cross-validating model 3

training set	train error	testing set	test error
partition 1+2		partition 3	
	0		5/10
partition 1+3			

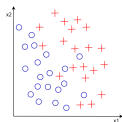


0

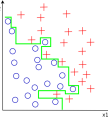
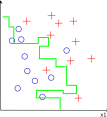
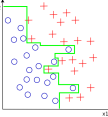


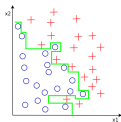
5/10

partition 1+3

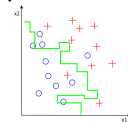


Cross-validating model 3

training set	train error	testing set	test error
partition 1+2		partition 3	
	0		5/10
partition 1+3			
	0		

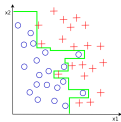


0



5/10

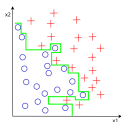
partition 1+3



0

Cross-validating model 3

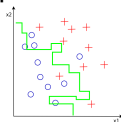
training set
partition 1+2



0

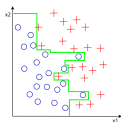
train error

testing set
partition 3



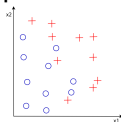
5/10

partition 1+3



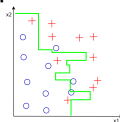
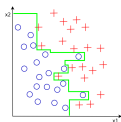
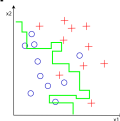
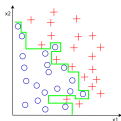
0

partition 2

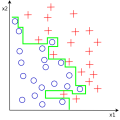
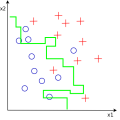
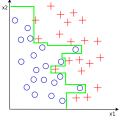
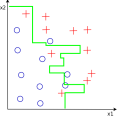
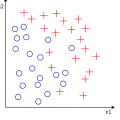


Cross-validating model 3

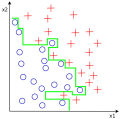
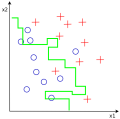
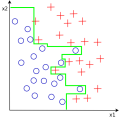
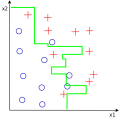
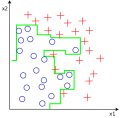
training set	train error	testing set	test error
partition 1+2	0	partition 3	5/10
partition 1+3	0	partition 2	5/10



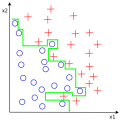
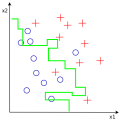
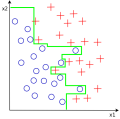
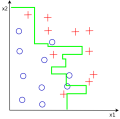
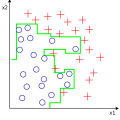
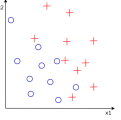
Cross-validating model 3

training set	train error	testing set	test error
partition 1+2 	0	partition 3 	5/10
partition 1+3 	0	partition 2 	5/10
partition 2+3 			

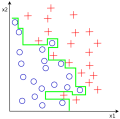
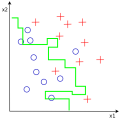
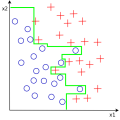
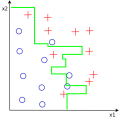
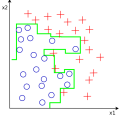
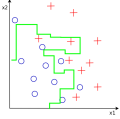
Cross-validating model 3

training set	train error	testing set	test error
partition 1+2 	0	partition 3 	5/10
partition 1+3 	0	partition 2 	5/10
partition 2+3 	0		

Cross-validating model 3

training set	train error	testing set	test error
partition 1+2 	0	partition 3 	5/10
partition 1+3 	0	partition 2 	5/10
partition 2+3 	0	partition 1 	

Cross-validating model 3

training set	train error	testing set	test error
partition 1+2 	0	partition 3 	5/10
partition 1+3 	0	partition 2 	5/10
partition 2+3 	0	partition 1 	6/10

Cross-validating model 3

training set	train error	testing set	test error	mean test error
partition 1+2		partition 3		
	0		5/10	
partition 1+3		partition 2		
	0		5/10	16/30
partition 2+3		partition 1		
	0		6/10	

Cross-validating models

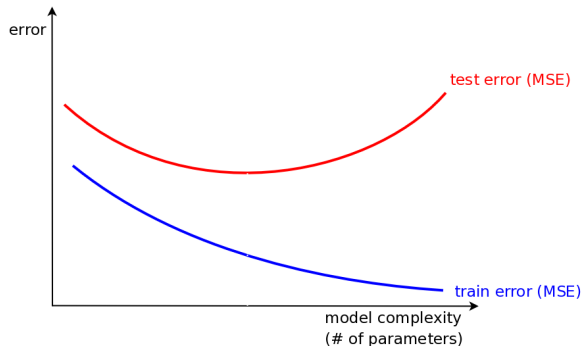
Summary

model	1	2	3
# of parameters	2	4	72
mean train error	15/60	12/60	0
mean test error	11/30	6/30	16/30

Cross-validating models

Summary

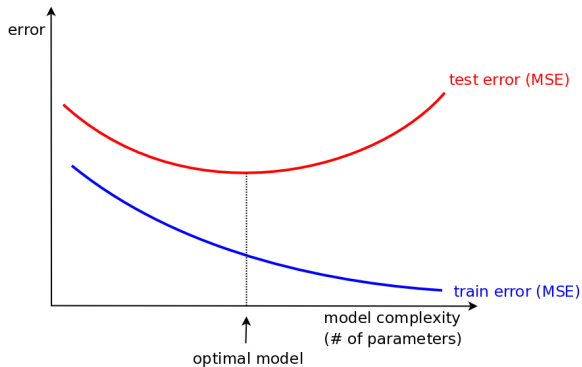
model	1	2	3
# of parameters	2	4	72
mean train error	15/60	12/60	0
mean test error	11/30	6/30	16/30



Cross-validating models

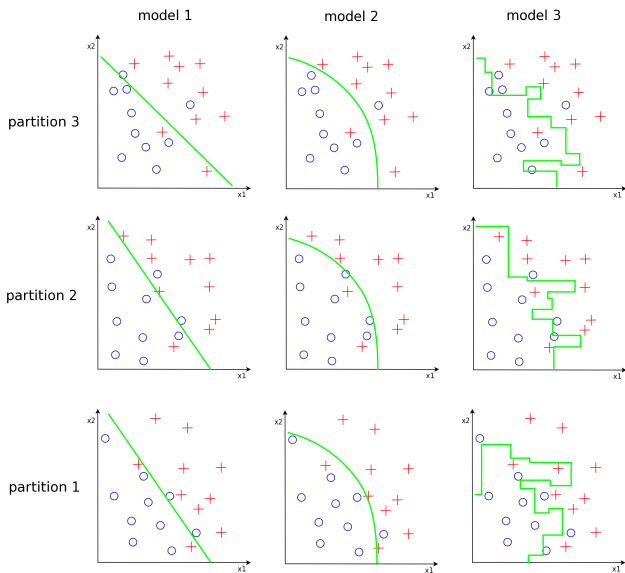
Summary

model	1	2	3
# of parameters	2	4	72
mean train error	15/60	12/60	0
mean test error	11/30	6/30	16/30



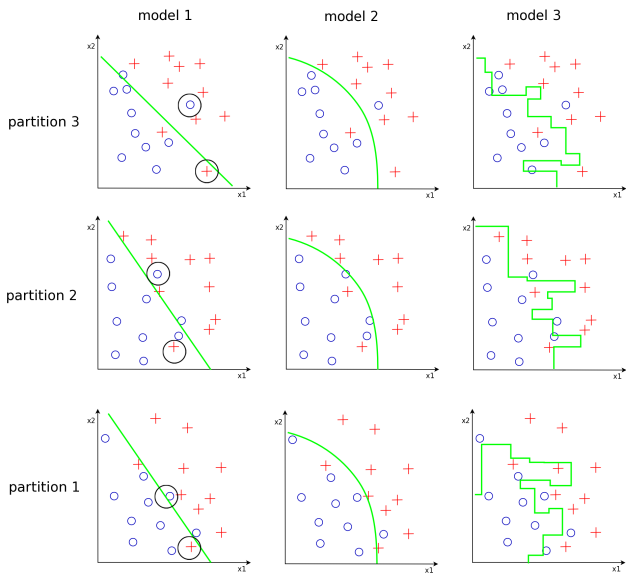
Model comparison

Test set performance



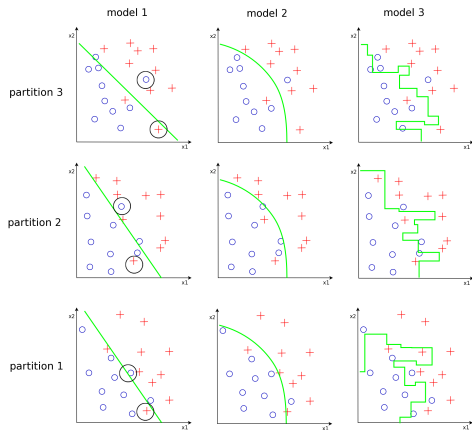
Model comparison

Test set performance



Model comparison

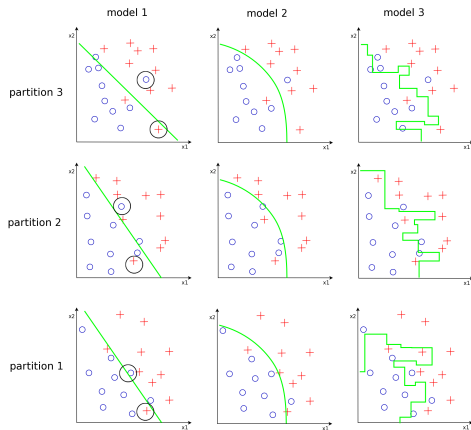
Test set performance



- model 1 is biased

Model comparison

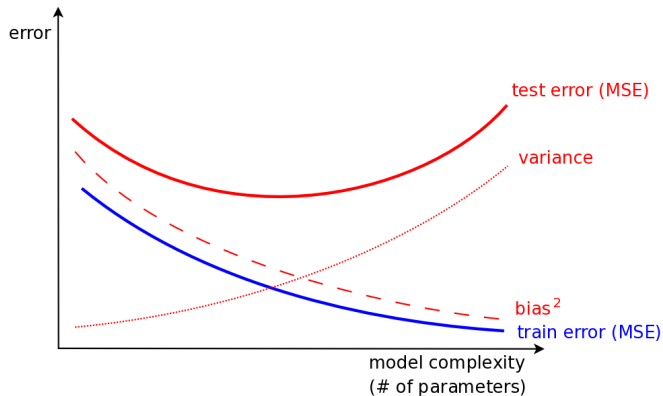
Test set performance



- model 1 is biased
- model 3 is highly variable and overfits the data

Cross-validation: summary

- $MSE = \text{bias}^2 + \text{variance}$



- high bias for low-complexity models
- high variance for high-complexity models

Cross-validation: conclusion

- CV can be seen as regularization technique
 - helps to overcome overfitting
 - note some models regularize naturally
 - e.g. nonparametric Bayesian models - see Maneesh's lectures
- CV can be seen as model selection technique