Model Selection

Machine Learning CSx824/ECEx242 Bert Huang Virginia Tech

Model Complexity

- Overfitting and underfitting
- Generalization error
- Validation for model selection





Underfitting

ML Algorithm 1

- Low dimensional
- Heavily regularized
- Bad modeling assumptions



- High dimensional or non-parametric
- Weakly regularized
- Not enough modeling assumptions
- Not enough data













Nearest-Neighbor Classifiers

0	\diamond	Ø	Ö	0
1		ł	4	l
Q	2	2_	2	2
3	3	3	3	3
ł	4	4	47	4
5	5	5	5	5
6	6	6	6	6
7	7	Ţ	7	フ
8	8	8	3	К
9	9	9	q	9

classifier = {

0:0, ****: 0, **2**:0, 0:0, **()**: 0, **ℓ**:1, : 1,

. . .

100% training accuracy!



53% testing accuracy...





Held-out Validation

0	\diamond	Ø	0	Ð
1		ł	4	l
2	2	2_	2	2.
3	3	3	3	3
벽	4	4	47	4
5	5	5	5	5
6	6	6	6	6
7	7	Ż	7	フ
8	8	8	3	К
4	9	9	q	9

Held-out Validation

1 1	0	\diamond	Ø	0	
2 2 2 3 Simple 3 3 3 3 3 3 4 4 4 4 4 Medium 5 5 5 5 5 Complex 6 6 6 6 Complex Super Complex 8 8 8 8 8 8 8	1		ł	4	
3 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 6 0	2	2	2_	2	Simple
44 44 44 Medium 55 55 55 55 60 66 66 Complex 7 7 77 77 8 8 8 8 8 9 9 7 77	3	3	3	3	
5 5 5 5 6 Complex 6 6 6 6 Complex 7 7 7 7 7 8 8 8 8 8 8 9 9 9 9 9 9	놱	4	4	47	Medium
6 6 Complex 7 7 7 7 7 8 8 8 3 Super Complex 9 9 9 4 1	5	5	5	5	
7 7 7 7 8 8 8 3 9 9 9 9 9	6	6	6	6	Compley
8 8 8 8 8 8 8 9 10 10 10 10 <td>7</td> <td>÷</td> <td>۲,</td> <td>7</td> <td></td>	7	÷	۲,	7	
9990000m	8	8	8	3	Super Com
	4	9	7	q	

training data

	Accuracy on training data	Accuracy on validation data
	0.91	0.83
	0.95	0.88
	0.99	0.79
X	1.0	0.54





\diamond	Ø	0	Ð
	ł	1	l
ູ	2_	2	2
3	3	3	3
4	4	47	4
5	5	5	5
6	6	6	6
7	7	7	フ
8	8	3	К
9	7	٩	9

Fold 1

training data



0	0	0	0
1	ł	ź	Ţ
2	2_	2	2.
3	3	3	3
Ł	4	47	4
5	5	5	5
6	4	6	6
7	Ţ	7	フ
8	8	3	С
4	7	q	9

training data

Fold 2



0	\diamond	0	0
1		4	
2	ູ	2	2
3	3	3	3
4	4	4	4
5	5	5	5
6	6	6	6
7	7	7	フ
8	8	3	К
9	9	٩	9

training data

Fold 3



0	\diamond	Ø	0
1		ł	l
2	a.	2_	2
3	3	3	3
4	4	4	4
5	5	5	5
6	6	6	6
7	7	Ţ	フ
8	8	8	К
9	9	7	9

training data

Fold 4



0	\diamond	Ø	0
1		ł	4
2	ູ	2_	2
3	3	3	3
4	4	4	47
5	5	5	5
6	6	6	6
7	7	7	7
8	8	8	3
9	9	7	٩

training data





	0	Ø	Ō	0
1		ł	4	l
2	2	2_	2	2.
3	3	3	3	3
ł	4	4	4	47
5	5	5	5	5
6	6	6	6	6
7	7	7	7	7
8	8	8	3	К
4	9	9	۹	9

training data



0		Ø	0	Ð
1		ł	ł	l
2	R	2_	2	2.
3	3	3	3	3
4	4	4	47	4
5	5	5	5	5
6	6	4	6	6
7	7	7	7	フ
8	8	8	3	С
9	9	9	٩	9

training data



0	\diamond		0	Ð
1		1	4	l
2	R	2_	2	2
3	3	3	3	3
ŁĮ	4	4	47	4
5	5	5	5	5
6	6	4	6	6
7	7	7	7	フ
8	8	8	3	С
9	9	7	q	9

training data



0	\diamond	Ø	0	
l		ł	1	Ш
2	ູ	2_	2	2
3	3	3	3	3
4	4	4	4	4
5	5	5	5	5
6	6	6	6	6
7	7	7	7	フ
8	8	8	3	К
9	9	7	q	9

training data



0	\diamond	Ø	0	0
		ł	4	l
d	ູ	2_	2	2
3	3	3	3	3
ł	4	4	49	47
5	5	5	5	5
6	6	4	6	6
7	7	7	7	7
8	8	8	3	С
4	9	7	4	9

training data



0	\diamond	Ø	Ö	0
1		ł	1	
2	ູ	2_	2	2
3	3	3	3	3
4	4	4	4	47
5	5	5	5	5
6	6	4	6	6
7	?	7	7	7
8	8	8	3	К
9	9	7	٩	9

training data



0	\diamond	0	0	0
1		ł	4	Ţ
	2	2_	2	2
3	3	3	3	3
ł	4	4	4	47
5	5	5	5	5
6	6	6	6	6
7	7	7	7	フ
8	8	8	3	К
4	9	7	q	9

training data



How Many Folds?

- What are the pros and cons of leave-one-out cross-validation?
- We usually train on N-1 folds and test on 1 fold. What are pros and cons of doing the inverse: train on 1 fold and test on N-1 folds?



Training





How Many Folds?

- What are the pros and cons of leave-one-out cross-validation?



• We usually train on N-1 folds and test on 1 fold. What are pros and cons of doing the inverse: train on 1 fold and test on N-1 folds?



Testing

Testing versus Validation

- Best practice for experiments:
 - Hold out test set completely hidden from training

 - Evaluate on held-out test data

Use validation on training data for model (or parameter) selection

Scenarios

- Mystery ML algorithm with single complexity parameter
- Scenario 1: cross-validation scores are extremely erratic

 Scenario 2: cross-validation score is very uniform



Model Selection via Validation

- Measure performance on **held-out** training data
 - Simulate testing environment
- Rotate **folds** of held-out subsets
- Can even hold out one at a time: leave-one-out validation
- Use (cross) validation performance to tune extra parameters