

1  
point

1. (True/False) A linear classifier assigns the predicted class based on the sign of  $\text{Score}(x) = \mathbf{w}^T h(\mathbf{x})$ .
- ☒ True
- ☐ False
- 

1  
point

2. (True/False) For a conditional probability distribution over  $y|\mathbf{x}$ , where  $y$  takes on two values (+1, -1, i.e. good review, bad review)  $P(y = +1|\mathbf{x}) + P(y = -1|\mathbf{x}) = 1$ .
- ☒ True
- ☐ False
- 

1  
point

3. Which function does logistic regression use to "squeeze" the real line to  $[0, 1]$ ?
- ☒ Logistic function
- ☐ Absolute value function
- ☐ Zero function
- 

1  
point

4. If  $\text{Score}(x) = \mathbf{w}^T h(\mathbf{x}) > 0$ , which of the following is true about  $P(y = +1|\mathbf{x})$ ?
- ☐  $P(y = +1 | x) \leq 0.5$
- ☒  $P(y = +1 | x) > 0.5$
- ☐ Can't say anything about  $P(y = +1 | x)$
- 

1  
point

5. Consider training a 1 vs. all multiclass classifier for the problem of digit recognition using logistic regression. There are 10 digits, thus there are 10 classes. How many logistic regression classifiers will we have to train?

10