

1.

(a) better

$n$  이 크고  $p$ 가 작은 경우에는 overfitting의 문제가 적고 데이터 간 관계에 대해 더 정확한 설명을 제공할 수 있다.

(b) worse

Overfitting의 문제가 발생할 수 있다.

(c) better

비선형적인 관계에서는 유연한 방법이 더 정확한 설명을 제공한다.

(d) worse

Error terms가 큰 경우에 overfitting의 문제가 발생할 수 있다.

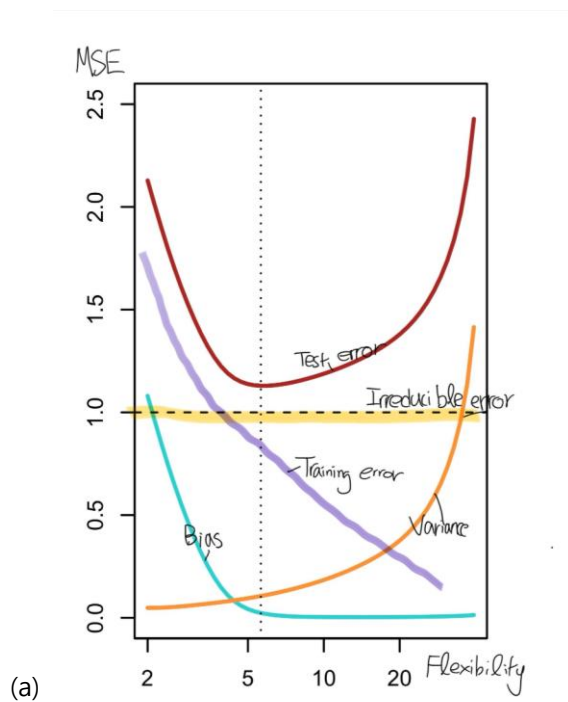
2.

(a) classification, inference

(b) classification, prediction

(c) regression, prediction

3.



(b)

Bias: 유연성이 증가할수록 편향은 감소한다.

Variance: 유연성이 증가할수록 분산은 증가한다.

Training error: 유연성이 증가할수록 training error는 감소한다.

Test error: 유연성이 증가함에 따라 bias의 감소로 줄어들었다가 variance의 증가로 다시 증가한다.

Irreducible error: 유연성과 무관하므로 불변한다.

5.

Flexibility는 Interpretability와 반비례 관계에 있는데, 추론의 경우 과도하게 유연한 기법은 X와 Y의 관계를 이해하기 어렵게 할 수 있다. 비선형적인 관계에서는 유연성이 높은 기법이 더 정확한 예측을 하게 할 수 있는데, 예측의 경우에도 더 유연한 기법이 더 부정확한 설명을 제공하는 경우가 있는데, 유연성이 높을수록 bias는 더 낮지만 variance는 높아지기 때문에 3번에서의 그래프에서처럼 극단적인 유연성에서는 큰 MSE 값을 가질 수 있다.

6.

Parametric:  $f$ 를 추정하는 것이 더 단순하다. 그러나 선택한 모델이  $f$ 의 실제 모양과 너무 다르면 정확도가 떨어진다.

Non-parametric:  $f$ 의 함수 형태에 대한 가정을 하지 않으므로 선택한 모델과  $f$ 의 실제 모양 간 괴리가 생기는 문제를 피할 수 있다. 그러나 정확성을 위해서는 parametric methods보다 더욱 많은 관측치를 필요로 한다.

7.

(a)

Obs

1.  $(3^2)^{0.5}=3$
2.  $(2^2)^{0.5}=2$
3.  $(1+9)^{0.5}=(10)^{0.5}=3.16$
4.  $(1+4)^{0.5}=5^{0.5}=2.23$
5.  $(1+1)^{0.5}=2^{0.5}=1.41$
6.  $(1+1+1)^{0.5}=3^{0.5}=1.73$

(b)  $K=1$

$$P(Y=\text{Red}|X=x_0)=1/1 \cdot I(y_5=\text{Red})=0$$

$$P(Y=\text{Green}|X=x_0)=1/1 \cdot I(y_5=\text{Green})=1$$

Green

(c)

$$P(Y=\text{Red}|X=x_0)=1/3(1+0+1)=2/3$$

$$P(Y=\text{Green}|X=x_0)=1/3(0+1+0)=1/3$$

(d) small

$K$ 값이 작을수록 더 유연하고 비선형적인 경계를 형성한다.