Training error ->

1. RMSE calculado no training set
2. Diminui com o aumento da complexidade do modelo
3. Overfitting, é viciado por ser calculado sobre os dados de treino (a não ser que os dados de treino representem TUDO no mundo)

Generalization (true) error ->

1. Considera a probabilidade de ocorrência das observações
2. Representado com o gráfico azul/branco
3. Diminui até certo ponto, mas depois aumenta com o aumento da complexidade do modelo
4. Não é possível computar este erro

Test error ->

1. Pode ser computado
2. RMSE calculado no test set
3. Coeficientes ainda foram calculados no training set
4. Semelhante ao generalization error, mas c/ mais ruído, em relação à complexidade do modelo

Overfitting ->

1. Acontece quando se tem um modelo para o qual:
   1. Training error (w1) < Training error (w2)
   2. True error (w1) > True error (w2)
   3. Vide curva de erro x complexidade

Noise ->

1. É inerente aos dados
2. Não pode ser controlado
3. Erro irredutível -> não pode ser modificado escolhendo um modelo ou procedimento diferentes

Bias ->

1. O quão bem um modelo pode contemplar a verdadeira relação entre x e y
2. Bias(x) = fw(true)(x) - fw(- average fit)(x)
3. O modelo é flexível o suficiente para capturar fw(true), a verdadeira natureza entre x e y? Se não, erro nas previsões
4. Diminui de acordo com o aumento da complexidade do modelo
5. Não pode ser computada pois depende da true function

Variance ->

1. O quão diferente pode ser meu modelo, para diferentes conjuntos de dados, entre um modelo e o modelo médio?
2. Aumenta de acordo com o aumento da complexidade do modelo
3. Não pode ser computada pois depende da true function

Bias-variance tradeoff ->

1. Ponto em que a complexidade do modelo é ótima para bias e variance
2. If the features of Model 1 are a strict subset of those in Model 2, the TRAINING error of the two models can **never** be the same.

False

1. If the features of Model 1 are a strict subset of those in Model 2, which model will USUALLY have lowest TRAINING error?

Model 2

1. If the features of Model 1 are a strict subset of those in Model 2. which model will USUALLY have lowest TEST error?

It's impossible to tell with only this information

1. If the features of Model 1 are a strict subset of those in Model 2, which model will USUALLY have lower BIAS?

Model 2

1. Which of the following plots of model complexity vs. RSS is most likely from TRAINING data (for a fixed data set)?

c

1. Which of the following plots of model complexity vs. RSS is most likely from TEST data (for a fixed data set)?

a

1. It is **always** optimal to add more features to a regression model.

false

1. A simple model with few parameters is most likely to suffer from:

High Bias

1. A complex model with many parameters is most likely to suffer from:

High Variance

1. A model with many parameters that fits training data very well but does poorly on test data is considered to be

Overfitted

1. A common process for selecting a parameter like the optimal polynomial degree is:

Minimizing validation error

1. Selecting model complexity on test data (choose all that apply):

Allows you to avoid issues of overfitting to training data

**Provides an overly optimistic assessment of performance of the resulting model**

Is computationally inefficient

Should never be done

Allows you to avoid issues of overfitting to training data

**Provides an overly optimistic assessment of performance of the resulting model**

Is computationally inefficient

**Should never be done**

1. Which of the following statements is true (select all that apply): For a **fixed model complexity**, in the limit of an infinite amount of training data,

The noise goes to 0

Bias goes to 0

Variance goes to 0

**Training error goes to 0**

Generalization error goes to 0

Nenhum

The noise goes to 0

Bias goes to 0

**Variance goes to 0**

Training error goes to 0

Generalization error goes to 0