

# DM-Spring-2020-Q6-Grade

78.95% (15/19)

- 
- ✓ 1. KNN is
- ☒ A data-driven method
  - ☐ B model-driven method
  - ☐ C I do not know
- ✓ 2. The dependent variable of the classification is
- ☒ A categorical
  - ☐ B numeric
  - ☐ C I do not know
- ✓ 3. KNN can be used for regression
- ☒ A Yes
  - ☐ B No
  - ☐ C I do not know
- ✓ 4. In the case of KNN classification we use
- ☐ A average of outcomes
  - ☒ B majority voting scheme
  - ☐ C I do not know
- ✓ 5. Which of these errors will increase constantly by increasing k?
- ☒ A train error
  - ☐ B test error
  - ☐ C both
  - ☐ D I do not know

✓ 6. This function can be used to perform KNN classification in R

- ☒ A knn()
- ☐ B k\_nn()
- ☐ C knnreg()
- ☐ D knearneib()
- ☐ E I do not know

✓ 7. With the increase of  $k$ , the decision boundary will be

- ☒ A simplified
- ☐ B more complex
- ☐ C I do not know
- ☐ D unchanged

✓ 8. KNN algorithm is sensitive to outliers

- ☒ A True
- ☐ B False
- ☐ C I do not know

✓ 9. KNN

- ☒ A is a supervised learning algorithm.
- ☐ B is an unsupervised learning algorithm.
- ☐ C I do not know

✓ 10. In the case of small  $k$  we have

- ☒ A overfitting
- ☐ B underfitting
- ☐ C it depends on the situation
- ☐ D I do not know

✓ 11. Why do we need scaling in KNN?

- ☐ A to avoid overfitting
- ☐ B to avoid underfitting
- ☒ C to have "equal" weights for variables
- ☐ D I do not know

- ✓ 12. Let  $k = n$ , ( $n$  - number of observations), K-NN is same as
- ☐ A random guessing
  - ☒ B everything will be classified as the most probable class (in total)
  - ☐ C everything will be classified as the least probable class (in total)
  - ☐ D I do not know
- ✓ 13. This function can be used to perform K-NN regression in R
- ☒ A knn.reg
  - ☐ B knnforreg
  - ☐ C regknn
  - ☐ D knnforregression
  - ☐ E I do not know
- ✗ 14. Do you need to worry about scaling with one explanatory variable?
- ☐ A No
  - ☒ B Yes
  - ☐ C I do not know
- ✓ 15.  $n$  - the number of observation,  $m$  - the number of explanatory variables When  $n=k$ ,  $m=1$ , the decision boundary for regression is
- ☒ A a line
  - ☐ B a stepwise constant function
  - ☐ C a stepwise quadratic function
  - ☐ D I do not know
- ✗ 16. Which of these algorithms can be used to fill the missing values
- ☐ A KNN for regression
  - ☒ B KNN for classification
  - ☐ C both
  - ☐ D I do not know



**17.** Which one is better: KNN regression or Linear regression?

- ☐ A KNN outperform LR if the parametric form that has been selected is close to the true linear form
- ☒ B LR outperform KNN if the parametric form that has been selected is close to the true linear form
- ☐ C KNN will always outperform the LR
- ☐ D I do not know



**18.** Which one is the Disadvantage of KNN?

- ☒ A required assumptions
- ☐ B cannot be applied for regression
- ☐ C difficult to perform
- ☐ D the problem of high dimensional data
- ☐ E I do not know



**19.** The best k for train set equals to

- ☐ A 1
- ☒ B 2
- ☐ C 0
- ☐ D I do not know