UNIVERSITY OF EDINBURGH
Business School

Predictive Analytics and Modelling of Data

CMSE11428 (2020-2021)

Dr Xuefei Lu
The University of Edinburgh Business School

Quiz

Based on 8 - Activity_6_performing_your_own_PCA.ipynb + absent.csv

Q1. You have performed PCA and can now answer what variables, or better, what linear combinations of variables are explaining the variance in the dataset well. What are the 5 most important variables in the first principle component?

- A. Reason for absence, Transportation expense, Distance from Residence to Work, Service time, Height
- B. Reason for absence, time, Age, Weight, Height
- C. Reason for absence, Transportation expense, Distance from Residence to Work, Service time, Weight
- **D. Transportation expense, Distance from Residence to Work, Service time, Age, Weight**

Quiz

Based on 8 - Activity_6_performing_your_own_PCA.ipynb + absent.csv

Q2. What are the 5 most important variables in the second principle component?

- **A. Reason for absence,Month of absence,Distance from Residence to Work,Work load Average/day,Height**
- B. Reason for absence,Month of absence,Distance from Residence to Work,Work load Average/day,Absenteeism time in hours
- C. Distance from Residence to Work,Work load Average/day,Weight,Height,Absenteeism time in hours
- D. Reason for absence,Month of absence,Distance from Residence to Work,Work load Average/day,Weight

Quiz

Based on 8 - Activity_6_performing_your_own_PCA.ipynb + absent.csv

Q3. Based on the two previous questions, which two variables do you think might be worthwhile for inclusion in the final model?

- A. Reason for absence,Distance from Residence to Work
- B. Transportation expense, height
- **C. Transportation expense,Work load Average/day**
- D. Work load Average/day,Weight

Quiz

Q1. What metrics are suitable for regression, rather than classification?

- A. RMSE, recall
- B. Recall and precision
- C. RMSE, MAPE
- D. MAPE, recall

Quiz

Q1. What metrics are suitable for regression, rather than classification?

- A. RMSE, recall
- B. Recall and precision
- **C. RMSE, MAPE**
- D. MAPE, recall

Quiz

Q2. What metric is more suitable to evaluate a skewed dataset?

- A. recall
- B. precision
- C. F-score
- D. AUC

Quiz

Q2. What metric is more suitable to evaluate a skewed dataset?

- A. recall
- B. precision
- C. F-score
- **D. AUC**

Quiz

03. Principle component analysis is what type of feature analysis technique?

- A. Filter method
- B. Wrapper method
- C. Embedded method
- D. Feature creation method

Quiz

03. Principle component analysis is what type of feature analysis technique?

- A. Filter method
- B. Wrapper method
- C. Embedded method
- **D. Feature creation method**

Quiz

Q4. Mutual information and Chi-squared values can be used to rank features. Therefore, they are:

- A. Filter method
- B. Wrapper method
- C. Embedded method
- D. Feature creation method

Quiz

Q4. Mutual information and Chi-squared values can be used to rank features. Therefore, they are:

- **A. Filter method**
- B. Wrapper method
- C. Embedded method
- D. Feature creation method