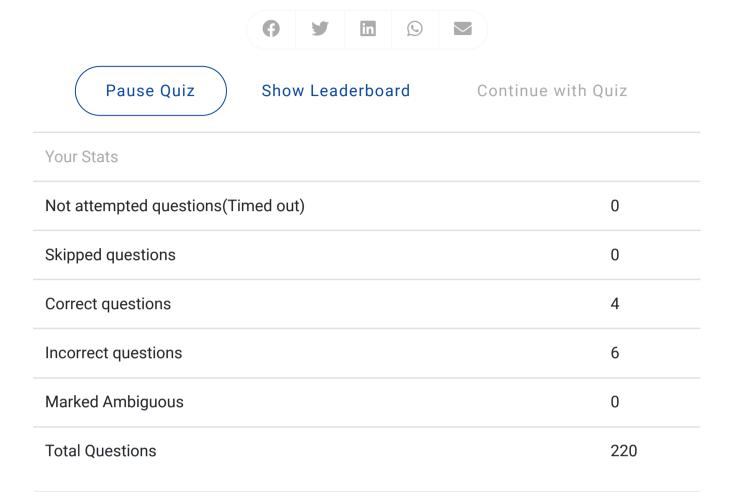
8/9/2019 Datamin | Review

# Review your last 10 questions

3 days, 1 hour remaining until quiz ends.



#### # 1/10

A startup is working on the recommendation engine for their platform. They first collected the ratings of different services and analysed. Which of the following results make the use of Latent Factor based model significant?

- Low correlation between rating of services
- High level of sparsity
- Low level of sparsity(approximately 0.1%)
- · None of the above



#### # 2/10

Kirk is working on a binary classification problem, which has a total of 1000 datapoints in the training set. Further it is given that this dataset has a total of 950 datapoints with a 'positive' label. He doesn't want to use Laplace smoothing.

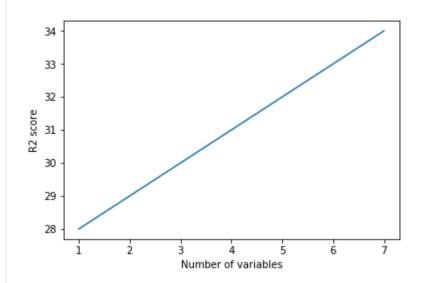
What will be the relation between the priors of both classes?

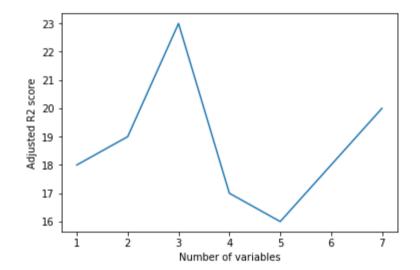
- Prior of positive label > prior of negative label
- prior of negative label > prior of positive label
- prior of positive label = prior of negative label
- None of the above



#### #3/10

Mr. X was trying to make a regression model. To make a good model he used different number of features. He then plotted two graphs one for r2 and another for adjusted r2. Based on theses two graphs which of the following is true?





- · Variables added after third variable were irrelevant
- Third variable was significantly more important
- All features are equally important
- Both a and c
- Both a and b

Incorrect

## # 4/10

A researcher concludes from his analysis that a placebo cures AIDS. He analysed a sample of size 1000. What type of error is he making?

- Type 1
- Type 2
- No error
- Cannot be determined

Incorrect

#### # 5/10

There are 15 students in a class. They all secure the following marks in a test out of 10:

10,9,4,6,8,9,1,6,8,10,4,6,7,7,6

Find out the geometric mean of the following data:

- 6.04
- 6.64
- 6.01
- 6.4

Correct

#### # 6/10

Which of these is not a hyperparameter in decision trees?

- criterion
- max\_depth
- max\_features
- n\_estimators

Incorrect

## # 7/10

In a DBSCAN algorithm, initially three clusters are made which are clearly separable. Which of the following statements is true if the maximum distance between the two points in a cluster started to increase (assuming the minimum number of samples in a cluster is low enough)?

- The model will be more outlier sensitive
- The number of clusters will start increasing
- The number of clusters will start decreasing
- Cannot be determined

Incorrect

# 8/10

8/9/2019 Datamin | Review

Rahul downloaded Google`s pretrained word2vec trained on google news dataset. He now wants to load the model using gensim library.

Help him complete the code.

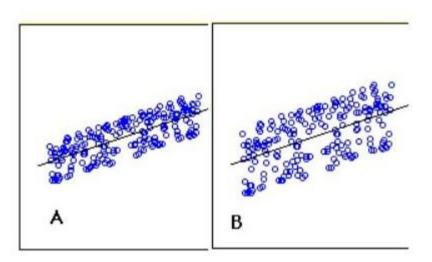
```
from gensim.models import Word2Vec

model = Word2Vec.___1___(`GoogleNews-vectors-negative300.bin`,
binary=True, norm_only=True)
```

- load\_word2vec\_format
- load\_word2vec
- load
- word2vec\_format



#### # 9/10



Which of the above graph has higher sum of residuals?

- A
- B
- sum of residuals of A and B are equal
- Need more data



### # 10/10

The sigmoid function is popular activation function. Which of the following is true for sigmoid function?

- It's maximum value is less than 0.5
- It's maximum value is 0.5
- It's maximum value is 1

4/5

8/9/2019

Datamin | Review • It's maximum value approaches infinity Correct i Suggested reading 7 Most Commonly Asked Questions On Correlation, Simple Linear Regression Pause Quiz **Show Leaderboard** Continue with Quiz **Analytics Vidhya Data Scientists** Companies About Us <u>Blog</u> Post Jobs Our Team <u>Hackathon</u> <u>Trainings</u> Hiring Hackathons <u>Careers</u> <u>Discussions</u> <u>Advertising</u> Contact us Apply Jobs **Download App** Visit us GET IT ON Google Play Sign up to our newsletter

your email address

© Copyright 2013-2019 Analytics Vidhya.

Privacy Policy Terms of Use Refund Policy