Week 1 quiz

Practice Quiz • 20 min

Review Learning Objectives

Submit your assignment

Receive grade

To Pass 80% or higher

Like **Dislike** **Report an issue**

Congratulations! You passed!

Grade received 92.86% **To pass** 80% or higher

Go to next item

1. You are training an NLP model to classify product reviews on a very large dataset, but training is taking a long time. How can you use feature engineering to reduce training time and possibly increase model performance? (choose all that apply)

1 / 1 point

- [ ] Randomly delete instances in the dataset to reduce the dataset size.
- [x] Filter irrelevant and redundant attributes, then retrain the model.

Correct
Correct! This is a feature selection approach that ensures that you are only keeping relevant attributes which is likely to reduce training time and possibly increase model accuracy..

- [x] Normalize or Standardize your data before training

Correct
Correct! Normalizing or standardizing data is the process of converting data to a common format/scale and can therefore reduce training time and sometimes increase model accuracy..

- [ ] None of the above

Try again

Your grade 92.86%

View Feedback

We keep your highest score

2. You perform correlation analysis on your feature set and discover that some features are highly correlated to each other. How can you take advantage of this information to improve your model's performance? (choose all that apply)

1 / 1 point

- [ ] Ignore this information as highly correlated features always improve model performance.
- [x] Combine the correlated features together.

Correct
Correct! Combining highly correlated features together can increase train speed and model performance as this will prevent the duplication of information.

- [ ] Apply different methods to increase the number of highly correlated features.
- [x] Eliminate one of the correlated features.

Correct
Correct! Features with high correlation have almost the same effect on the target variable and it therefore makes sense to remove one of them.

3. You have a task to train a text classifier on a customer product reviews dataset.. You decide to use the "star rating" to create 3 sentiments.

Rating 1 & 2 = Negative

Rating 3 = Neural

Rating 4 & 5 = Positive

What feature engineering method did you use in this scenario?

1 / 1 point

- ( ) Feature Selection
- (•) Feature Creation
- ( ) Feature Transformation
- ( ) All of the above

Correct
That's right! In feature creation we can combine existing features into new features or even create new attributes from existing ones. This is exactly what was done here as new sentiment features have been created from the "star rating" attribute.

4. BERT, which stands for Bidirectional Encoder Representations from Transformers, and Amazon SageMaker BlazingText are 2 popular natural language processing (NLP) algorithms. What are some characteristics of these algorithms?

1 / 1 point

- [ ] Unlike BERT, BlazingText cannot generate vectors for words encountered outside its vocabulary space i.e it does not support out-of-vocabulary words.
- [ ] They both take into account the word position when generating the embedding.