Practical Data Science Specialization

← Week 1
Practice Quiz • 30 min

✓ **Congratulations! You passed!**
TO PASS 70% or higher

Keep Learning

GRADE
100%

Week 1

TOTAL POINTS 6

1. Suppose you want to use Machine Learning to classify sentiments of product reviews. The model inputs will be the product reviews text and the output the predicted sentiment, positive or negative. In a typical ML workflow there are 3 main steps:

   I. Train and tune the machine learning model

   II. Deploy and monitor the trained model

   III. Prepare the data

   What is the correct ordering of these steps?

   ● III, I, II

   ○ III, II, I

   ○ I, II, III

   ○ I, III, II

   1 / 1 point

2. Data lakes are enterprise storage solutions which can host virtually any amount of data. Are there any restrictions on the data types a data lake can store?

   ○ Yes, data lakes only support structured data.

   ○ Yes, data lakes only support unstructured data.

   ○ Yes, data lakes only support unstructured and semi-structured data types.

   ● No, data lakes support all data types including structured relational data, semi-structured data, and unstructured data.

   1 / 1 point

   > ✓ **Correct**
   > Correct! A data lake stores data in its original format, without having to first structure the data.

3. Amazon Simple Storage Service (Amazon S3) is a public cloud object storage service that allows users to store and retrieve any amount of data at any time.

   Select the tool that you would use to register and discover data in Amazon S3.

   ○ AWS Data Wrangler

   ○ Amazon Athena

   ○ Amazon Redshift

   ● AWS Glue

   1 / 1 point

4. Amazon Athena provides great flexibility to run queries without adding any complexity to your project. Moreover, it is a very fast service and your queries return results in a matter of seconds, even on large datasets. Which of the following facts is **NOT true** about Amazon Athena?

`1 / 1 point`

○ No infrastructure is needed to set up Amazon Athena.

⦿ Amazon Athena does not require the AWS Glue Data Catalog to register and query S3 data.

○ It is an interactive query service.

○ It can be used to analyze data in Amazon S3 using standard SQL.

5. Within the AWS ecosystem, Data Wrangler is an agile service to load and unload data from data lakes and databases. What are other capabilities of this service? (Choose all that apply.)

`1 / 1 point`

☑ It connects pandas dataframes and other AWS services.

> ✓ **Correct**
> Correct! You can read more about AWS Data Wrangler here.

☑ It extends the power of the pandas library to AWS.

> ✓ **Correct**
> Great Job. You can read more about AWS Data Wrangler here.

☑ It can be used for loading and unloading data from data lakes and databases.

> ✓ **Correct**
> That's right! You can read more about AWS Data Wrangler here.

☐ It extends the power of the NumPy library to AWS.

6. You work as a Machine Learning engineer in a company and are asked to develop algorithms to solve 3 tasks:

`1 / 1 point`

Task 1: You have a large dataset of unstructured text information. You are asked to convert/summarize this information into reports.

Task 2: A business has experienced a huge surge in the number of customers. To improve customer support and engagement, you are asked to build a chatbot.

Task 3: To simplify paperwork and time, you are asked to automate an employee expense system by building an image scanning system for expense receipts.

Which of these will you treat as Natural Language Processing (NLP) problems?

○ Tasks 1 and 3

○ All 3 tasks

⦿ Tasks 1 and 2

○ Task 1 only

> ✓ **Correct**
> Correct! NLP is appropriate for text summarization task (Task 1) and an NLP-based Chatbot that uses text or sound can be used for customer support (Task 2).