Announcement

Close

Unit 3

Tue, Oct 22, 7:11AM CT | 0 Attachment(s)

Unit 3

Reading Task:

- Data Science Versus Big Data Versus Data Analytics
  - This article talks about the differences between data science, big data, and data analytics.
- The 4 Types of Data Analytics
  - This Web site explores 4 different types of data analytics and provides insight into how to apply them effectively.
- Machine Learning
  - This video explores the fundamentals of machine learning.
- Spark and Machine Learning
  - This video examines machine learning and how it can use streams of data.
- Scikit-Learn Tutorial
  - This tutorial provides an overview of Scikit-Learn (a popular Python machine learning library).
- From Machine Learning to Machine Reasoning: An Essay
  - This article describes the manipulation applicable to training systems to build reasoning capabilities.
- Mastering Machine Learning Algorithms: Expert Techniques to Implement Popular Machine Learning Algorithms and Fine-Tune Your Models
  - This article discusses various techniques used to implement machine learning algorithms.

Unit 3 - Discussion Board

Task Type: Discussion Board

Deliverable Length: See assignment details

Points Possible: 75

Description:

Primary Response is due by Thursday (11:59:59pm Central), Peer Responses are due by Saturday (11:59:59pm Central).

Primary Response: Within the Discussion Board area, write 300-500 words that respond to the following questions with your thoughts, ideas, and comments. This will be the foundation for future discussions by your classmates. Be substantive and clear, and use examples to reinforce your ideas.

For this Discussion Board, please complete the following:

Machine learning has been around for several decades, but it has recently exploded due to much more data being available, more powerful computing, and more tools that are easier to use. Machine learning is different from traditional programming in that it allows a data scientist to create a learning algorithm that identifies patterns and uses these for future predictions.

For this task, you will research the application of machine learning to streaming data.

An example in your daily life is an online gaming company that collects streaming data about player-game interactions and feeds the data into its gaming platform. It then analyzes the data in real-time and offers incentives and dynamic experiences to engage its players.

- Describe the differences between streaming data and at-rest data.
- What are other streaming data examples and potential algorithms for use in your daily life?
- Would you recommend stream or batch processing for processing the big data collected from online gaming? Explain your decision.

Responses to Other Students: Respond to at least 2 of your fellow classmates with at least a 100-200-word reply about their Primary Task Response regarding items you found to be compelling and enlightening. To help you with your discussion, please consider the following questions:

- What did you learn from your classmate's posting? What additional questions do you have after reading the posting?
- What clarification do you need regarding the posting?
- What differences or similarities do you see between your posting and other classmates' postings?

For assistance with your assignment, please use your text, Web resources, and all course materials.

Discussion Board Rubric

Expectation	Points Possible	Points Earned	Comments

Application of Learning Material Content: Post demonstrates understanding of Learning Material content.	20		
Application of Course Knowledge: Post contributes unique perspectives or insights gleaned from text/learning resources, or specified by assignment.	20		
DB Responses: Responds substantively to two posts. Responses encourage interaction in the Discussion Board and classroom community.	20		
Organization: Post presents information logically and is clearly relevant to discussion topic.	8		
Professional Language: Posts contain accurate grammar, spelling, and/or punctuation with few or no errors. **Any resources should be cited in APA format or style specified in the assignment.**	7		
Total Points	75		
Total Points Earned			

View a downloadable version of the Discussion Board Grading Rubric.

Course Objectives:

- Understand the differences between streaming data analytics and at-rest data analytics, as well as the foundation, techniques, and processes of the streaming data analytics.

Model Answer:

Unit 3 - Individual Project

Task Type: Individual Project

Deliverable Length: 3-5 pages (not including cover page and resource page)

Points Possible: 125

Description:

Assignment Details

Combining machine learning with streaming data can create an environment where predictions can be made in real-time based on tangible data. For instance, if a manufacturer were to embed an algorithm in an automobile computer control system, it could collect data that will help predict parts failures before they occur.

For this assignment, you will consume the data from this link and consume data locally on your PC in Python. Construct a machine learning algorithm on these data. The algorithm may be supervised or unsupervised.

The project deliverables include the following:

- Create an ingest job in Python that consumes the identified data.
- Construct an unsupervised or supervised machine learning algorithm to predict the numerical or categorical outcome based on the defined goal.
- Discuss the results.
- Discuss the solution design, interworkings, and how the algorithm is to be interpreted.
- Once complete, submit your assignment for grading in a Microsoft Word document.

Individual Project Rubric

The Individual Project (IP) Grading Rubric is a scoring tool that represents the performance expectations for the IP. This Individual Project Grading Rubric is divided into components that provide a clear description of what should be included within each component of the IP. It's the roadmap that can help you in the development of your IP.

Expectation	Points Possible	Points Earned	Comments
Identification: Identify the differences between streaming data analytics and at-rest data analytics, as well as the foundation, techniques, and processes of the streaming data analytics.	30		
Analysis: Analyze the preparation, collecting, storing, securing, analyzing, interpreting, processing, and reporting of streaming data.	30		
Python: Ingest the data from Twitter into Python.	30		
Organization: Assignment presents information logically and is clearly relevant to discussion topic.	25		
Professional Language: Assignment contains accurate grammar, spelling, and punctuation with few or no errors.	10		

Total Points	125		
Total Points Earned			

Reference

Blum, S. (2021). *Twitter stream*. PubNub. https://www.pubnub.com/developers/realtime-data-streams/twitter-stream/

Course Objectives:

- Understand the differences between streaming data analytics and at-rest data analytics, as well as the foundation, techniques, and processes of the streaming data analytics.

Model Answer:

Student answers will vary, but they should include the following:

- The solution shall ingest the Twitter data into Python.
- Python libraries should be used to construct an unsupervised algorithm (e.g., association rules, clustering, or dimension reduction) and a supervised algorithm (e.g., classification or regression).

Unit 3 - Live Chat Extra Credit Summary

Task Type: Live Chat

Deliverable Length: 200 words

Points Possible: 15

Description:

In this class, you have the opportunity to receive credit for attending Live Chats or reviewing the chat archive. This opportunity is available once per week. To receive these points, you can either:

1. Attend the Live Chat session.
   or

2. Review the archived chat session and submit a 200-word summary of the content. The archive review summary must meet the expectations described below.

- Summarize the content of the chat. Some ideas for what to include are a description of what information was covered and how you will apply this information to your academic or professional work.
- Label the document "Chat Credit" with the date the Live Chat was presented.
- A summary must be submitted within 1 week of the Live Chat to be considered for credit.

Your instructor has the discretion to determine whether to award the points. A summary must meet each of these guidelines. Points are awarded on an all-or-nothing basis.

To submit your summary, head to the assignment list and select Unit 3 - Live Chat Extra Credit Summary.

View the Individual Project Demo for guidance on how to submit your chat summary as an Individual Project via the Virtual Classroom or the CTU Mobile app.

Course Objectives:

- Understand the differences between streaming data analytics and at-rest data analytics and the foundation, techniques, and processes of streaming data analytics.

Model Answer:

Students are required to meet all of the expectations as outlined in the assignment.

Optional Live Session Talking Points:

- What is Analytics?
- What is Machine Learning?
- Interpreting the Results of a Machine Learning Algorithm

Attachments:

There are no attachments for this announcement.