Cloud Computing for Data Analysis

DSBA 6190-U90 (CRN 23148)

Instructor Information

Instructor	Email	Office Location & Hours
Colby T. Ford, Ph.D.	colby.ford@uncc.edu	PORTAL 354 (By Appt.)

General Information

Wednesdays 5:30 pm - 8:15 pm, Dubois Center 801

Description

Introduction to the basic principles of cloud computing for data intensive applications. Covers a broad range of technologies and solutions from data platform architecture to data analytics. Focuses on the scalable deployment of cloud resources and the integration between individual services. Topics covered will include the deployment and use of data architectures with data lakes and data warehouses, containerized applications, distributed computing using cluster technologies such as Apache Spark, machine learning and deep learning using scalable/GPU-based infrastructure.

Restriction(s): Familiarity with Python and/or R, SQL, Unix, Data Structures, Machine Learning Algorithms, and Statistics; good programming skills and a solid mathematical (linear algebra) background.

Learning Outcomes:

1. Understand the benefits of cloud-based architecture
2. Architect end-to-end solutions based on user/organizational requirements
3. Recognize the differences in data platform options on-premises versus in the cloud
4. Discuss the cloud and on-premises machine learning approaches and the benefits therein

Grading:

The final course grade will be determined by the student's total number of points earned in the class out of the total possible points.

Exercise	Points
Data Platform Lab	100
Distributed Computing Lab	150
Midterm Exam	200
Machine Learning Lab	150
DevOps Lab	100
Final Exam	300
Total	1000

Final Grading Ranges	
≥900/1000pts	A
800-899/1000pts	B
700-799/1000pts	C
<700/1000pts	D or Inc.
Academic Dishonesty	F

Tentative Course Schedule

Date	Section	Topic(s)	Lab
January 11th	Intro (VIRTUAL)	- Introduction to the course - Gitting Started/App Setup	
January 18th	Data Platform	- Unstructured Data Stores o Blob Storage o Data Lakes	Begin Data Platform Lab
January 25th	Data Platform	- Structured Data Stores o Databases o Data Warehouses - Azure Synapse and Data Factory	
February 1st	Data Platform / Distributed Computing	- Containerization (Docker) o Kubernetes	Data Platform Lab Due Begin Distributed Computing Lab
February 8th	Distributed Computing	- Apache Spark (Databricks)	
February 15th	Distributed Computing / Review	- Other Cluster Technologies - Intro to Machine Learning and Neural Networks - Review Data Platform Topics	Distributed Computing Lab Due
February 22nd	Midterm Exam		
March 1st	NO CLASS		
March 8th	Machine Learning	- ML Options in the Cloud - Cognitive Services - Intro to Azure Machine Learning	Begin Machine Learning Lab
March 15th	Machine Learning	- Regression, Classification, Clustering - Automated Machine Learning	
March 22nd	Machine Learning	- Cross-Validation and Tuning - Accelerating Training with Distribution and GPUs	
March 29th	Machine Learning	- MLOps and Model Deployment	Machine Learning Lab Due
April 5th	DevOps	- Arch. Deployment with Terraform	Begin DevOps Lab
April 12th	DevOps	- CI/CD Pipelines with GitHub Actions	
April 19th	Security and Compliance	- RBACs, ACLs, and AAD - Compliance Considerations	
April 26th	Review	- Review Distributed Computing Topics - Review DevOps Topics	DevOps Lab Due
May 10th	Final Exam		DUBOIS 801 5:00pm-7:30pm

Academic Integrity and Honesty:
Students are required to read and abide by the Code of Student Academic Integrity available from Dean of Students Office. This code forbids cheating, fabrication or falsification of information, multiple submissions of academic work, plagiarism (including viewing others work without instructor permission), abuse of academic materials, and complicity of academic dishonesty. Violations of the Code of Student Academic Integrity, including plagiarism, result in disciplinary action as provided by the Code.

Civility:
We are concerned with a positive learning experience. This course strives to create an inclusive academic climate in which the dignity of all individuals is respected and maintained. We value diversity that is beneficial to both employers and society at large. Students are encouraged to actively and appropriately share their views in class discussions.

Inclement Weather:
University Policy Statement #13 states the University is open unless the Chancellor announces that the University is closed. The inclement weather hotline number to call is 704-687-1900. In the event of inclement weather, check your e-mail, and Canvas. The instructor will post a message on Canvas, and through e-mail. The instructor will use their best judgment as to whether class should be held.

Illness:
If you are feeling unwell, have a fever, have symptoms of COVID-19 or another illness (flu, RSV, etc.), or have been in contact with someone with COVID-19 or other illness, you are not allowed to attend class.

Disability:
UNC Charlotte is committed to access to education. If you have a disability and need academic accommodations, please provide a letter of accommodation from Disability Services early in the semester. For more information on accommodations, contact the Office of Disability Services at 704-687-0040 or visit their office in Fretwell 230.

Withdrawal:
The University policy on Course Withdrawal allows students a limited number of opportunities available to withdraw from courses. There are financial and academic consequences that may result from course withdrawal. If a student is concerned about his / her ability to succeed in this course, it is important to make an appointment to speak with the instructor as soon as possible.

Syllabus Revision:
The instructor may modify the class schedule and syllabus throughout the semester. Changes will appear on Canvas. Students are responsible for refreshing their syllabus.

Email Communication:
Students are responsible for *all* announcements made in class and on the class online resources. Students should check the online class resources throughout the semester. The Instructor and Teaching Assistants send occasional e-mails with important information. We send this information to the student's UNC Charlotte e-mail address listed on Banner system. If a student is not checking his / her UNC Charlotte e-mail address (ex. userName@uncc.edu) please be sure to access this e-mail and check it regularly during this course.