Course Information

Course Number/Section: ITCS 6155-8155, DSBA 6155
Course Title: Spring 2022
Days & Times: Tuesdays Thursdays 8:30 to 9:45 Atkins 126

Instructor Contact Information

Instructor: Dr. Pamela Thompson
Email Address: plthomps@uncc.edu
Phone: 980-234-0042 (text, call)
Office Location: Woodward Hall
Office Hours: 7:30 – 8:30 pm Monday Wednesday
https://us06web.zoom.us/j/2694772139
and by individual appointment (virtual):

Course Prerequisites

- Graduate/Ph.D. student standing or permission of instructor.
- ITCS 6162 Knowledge Discovery in Databases
- Familiarity with programming language (Python, Java or C++)

Course Descriptions

ITCS 6155 /8155 Knowledge Based Systems Credit 3(3-0)
Prerequisite: ITCS 6162 Knowledge Discovery in Databases or permission of instructor. Knowledge systems; knowledge discovery; association rules; action rules, hierarchical classifiers, cascade classifiers, query languages and their semantics; cooperative and collaborative systems; ontology and metadata; flexible query answering; chase algorithms and data sanitization methods; decision support systems in medicine; and automatic indexing of music.

About this Class: Knowledge Based Systems will be explored and implemented in a cloud environment using Big Data. The topic will be covered from the standpoint of a data engineer focusing on all aspects of the KBS including acquisition and preparation of data, statistical analysis, predictions based on algorithms that build models (classifiers, neural networks, association rules) for knowledge and the evaluation of the effectiveness of the recommendations. The data engineer goes from building the statistical and machine learning models to automating them. The resulting Knowledge Based System is secure, reliable, fault-tolerant, scalable and efficient.

We will be using a cloud based environment such as Google Cloud Platform to build a sophisticated statistical and machine knowledge based system.
A more traditional view of the knowledge based system is a system built from rules that come from problem solving expertise from human and domain experts. The main stages in the development of the knowledge based agent are similar to the knowledge discovery process: understanding the domain, modeling problem solving in the domain, developing the ontology and learning the reasoning rules, then testing the agent. A wide variety of domains require evidence-based reasoning like this, including intelligence analysis, cybersecurity, law, forensics, medicine, and education. We will take some time in class to look at knowledge based systems in the traditional sense also.

Student Learning Objectives
The objectives of this course are to design and develop a knowledge based system by
- Applying statistical and machine learning methods to build a data driven knowledge based recommender system on Google Cloud Platform.
- Create a high-performing prediction model with TensorFlow.
- Build a logistic regression machine learning model with Spark.
- Create a Bayesian model on a Cloud Dataproc cluster.
- Build a real-time analysis pipeline to carry out streaming analytics.
- Build a knowledge based reasoning agent with Disciple-EBR.
- Utilize evidence-based reasoning that learns complex problem-solving expertise directly from human experts, support experts, and nonexperts in problem solving and decision making in order to build the knowledge base.
- Be able to articulate the differences between knowledge based systems built on machine learning and knowledge based systems relying on rules and other forms of data for inference.
Student Learning Outcomes
An ability to design, implement, and evaluate a knowledge based system utilizing an end to end data pipeline using statistical and machine learning methods and tools on a cloud platform (Google Cloud Platform or other).

Assessment Criteria:

- Interactive Tutorials: Qwiklabs, Datacamp 30 points
- Applied Exercises 40 points
- Project 30 points
- PhD Students will also be required to complete a review of a scholarly research paper (2020 or later) related to Knowledge Based Systems (evaluation of a minimum of three recent articles in a formal paper – due by last class, end of day)

Late work: Late work is only accepted with prior approval. Work that is received after the due date without prior approval is automatically deducted 15 points. Work over a week late will not be accepted.

Required Textbook(s), Software:

Data Science on the Google Cloud Platform, O’Reilly, Valliappa Lakshmanan, 2022.

Google provides the class with a grant in order to take advantage of Google Cloud Platform and Qwiklabs for class. Students receive credits in $50 increments and are taught how to use cloud resources effectively and with a budget.
### Tentative Schedule/Assignments & Academic Calendar

NOTE: The following class schedule and deadlines are subject to change at the discretion of the instructor.

<table>
<thead>
<tr>
<th>Date</th>
<th>Topic</th>
<th>Assignment</th>
<th>HW, Project</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>Important Dates:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>First Class: Thursday 1/10</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Spring Break: March 7 - 11</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Last Class: Tuesday May 3rd</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Final Exam: Project Presentations Tuesday May 10th 8 to 10:30 am</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LESSON</td>
<td>NOTES</td>
<td>Note: Work is assigned to categories: Tutorials, Applied Exercises, Project</td>
<td></td>
</tr>
<tr>
<td>WEEK 1</td>
<td>Review Syllabus, Sign up for accounts</td>
<td></td>
<td>Qwiklab 1: GCP Essentials Due end of day Monday 1/24</td>
</tr>
<tr>
<td></td>
<td>Start first Qwiklab Quest, Complete Student Introduction</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WEEK 2</td>
<td>1: Making Better Decisions</td>
<td>Qwiklab 1: GCP Essentials Due end of day Monday 1/24</td>
<td>Exercise 1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WEEK 3</td>
<td>2: Ingesting Data in the Cloud</td>
<td>Qwiklab 2: Baseline: Data, ML, AI</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Qwiklab 3: Intro to SQL for BigQuery &amp; Cloud SQL</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Qwiklab 4: Ingesting Data Into the Cloud</td>
<td></td>
</tr>
</tbody>
</table>
| WEEK 4 | 3: Creating Compelling Dashboards | Qwiklab 5: Ingesting Data Into the Cloud Using Google App Engine  
Qwiklab 6: Loading Data into Google Cloud SQL  
Qwiklab 7: Visualizing Data with Google Data Studio |
| WEEK 5 | 3: Creating Compelling Dashboards | Exercise 2 |
| WEEK 6 | 4: Streaming Data: Publication and Ingest  
5: Interactive Data Exploration | Qwiklab 8: Processing Data with Google Cloud Dataflow  
Qwiklab 9: Visualize RT Geospatial Data w Google Data Studio  
Qwiklab 10: Loading Data into Big Query for EDA  
Exercise 3 |
| WEEK 7 | 4: Streaming Data: Publication and Ingest | Qwiklab 11: EDA Using AI Platform  
Exercise 4 |
| WEEK 8 | 5: Interactive Data Exploration | Qwiklab 12: Evaluating a Data Model  
Exercise 5  
Datacamp: PySpark |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>WEEK 9</td>
<td>March 7 - 11</td>
<td></td>
</tr>
</tbody>
</table>
| WEEK 10 | 6: Bayes Classifier on Cloud DataProc  
7: Machine Learning: Logistic Regression on Spark | Start Qwiklabs second Qwest:  
Datascience on GCP: Machine Learning  
Qwiklab 13: ML with Spark on Google Cloud Dataproc  
PROJECT ASSIGNMENT |
| WEEK 11 | 8: Time-Windowed Aggregate Features | Qwiklab 14: Processing Time Windowed Data  
Exercise 6 |
| WEEK 12 | 9: Machine Learning Classifier Using TensorFlow | Qwiklab 15: ML with Tensorflow  
Qwiklab 16: Distributed ML with Google Cloud ML |
| WEEK 13 | 10: Real Time Machine Learning | Qwiklab 17: Real Time ML with Google Cloud ML  
Exercise 7 |
| WEEK 14  | Rules Based Systems               | HW: Rules Based Systems: Watson Assistant – building a rules based expert system  
|          |                                 | Exercise 8                  |
| WEEK 15  | PROJECT WORK                     | PROJECT WORK                |
| WEEK 16  | PROJECT WORK                     | PROJECT WORK                |
| LAST DAY OF CLASS  
Tuesday May 3rd  
Exam: Tuesday May 10th 8 to 10:30 am  | Projects Due by Start of Final Exam  
All HW due by end of day Tuesday May 3rd  | All work due by end of day Tuesday May 3rd |
Course Format and Activities
This course will draw materials primarily from the required textbook and handouts/materials posted on the course website. Students will study the materials and complete all the course requirements. In order to properly address the assignments for this class, you will need to put in a considerable amount of time and energy. Please log on often to check for announcements, assignments, course documents, news forums, grades daily to stay informed. Students are expected to log in one or more times every 48 hours.

Reading:
The reading for this course will be taken from the textbook (primary) and a variety of other current sources. There will be assigned readings each week. Students must read the course materials and post any questions that you wish to be discussed on the forum.

Group discussion:
The most vital use of the Forums is to exchange ideas with other classmates. It is important that you check into the forums regularly. You are encouraged to ask questions regarding to the required readings, discuss the unit topics, share information and resources with classmates, and respond to problems posted by your classmates or instructor. You should read everyone’s posts and responses to the topics that interest you.

Project: (outlined in a separate assignment available in Canvas at midpoint for the course)

Submission of Work:
• Follow each assignment instruction; all work should be uploaded into the assignment link, or the Discussion board on Canvas or on Piazza as instructed. It is the students’ responsibility to keep his/her copies of all work submitted to the instructor. All work is to be turned in by the due date, no late work will be accepted without prior approval. Late work will automatically receive a 15 point deduction and will not be accepted after 1 week late.
Policy on Academic Integrity: The university policy 407, the Code of Student Academic Integrity, applies. This policy is available at http://legal.uncc.edu/policies/up-407.

Academic honesty is absolutely essential. Cheating, plagiarism or other academic misconduct will not be tolerated. If you are caught cheating, you will not pass this course and will be subject to any and all penalties specified in the code of Student Academic Integrity. If a student is found cheating, she or he will receive a ZERO for that assignment. If a student is found cheating a second time, she or he will receive an “F” for the course.

Examples of violation academic integrity include, but are not limited to:

- pretending that somebody else's work is yours so that you can get a higher grade than your own work merits
- falsifying data
- lying in order to extend a deadline or gain some other special advantage
- helping other people to do any of these things
- copying answers on tests
- using prohibited reference materials (such as notes or books) during an exam
- turning in papers that you have not written yourself or that you wrote for a different course
- quoting material without marking it as quoted and without attributing it to its source (or closely paraphrasing material without attributing it to its source)
- misrepresenting a medical or family emergency or other personal contingency in order to delay a scheduled exam or to get extra time on an assignment
pretending to have a disability you do not have (or exaggerating one you do have) in order to gain an unwarranted advantage unavailable to other students
modifying graded material and then resubmitting it to "correct the error in grading"

Rules Governing Students with Special Requirements
Students with disabilities which require accommodations should:
1. Register with the Office of Disability Support Services and 504 Compliance to provide documentation
2. Bring the necessary information indicating the need for accommodation and what type of accommodation is needed. This should be done during the first week of classes or as soon as the student receives the information. If the instructor is not notified in a timely manner, retroactive accommodations may not be provided.

Miscellaneous Requirements
1. All requests to change grading of any course work must be submitted in writing within a week after the grades are made available. Requests must be specific and explain why you feel your work deserves additional credit.
2. All requests about missing (or zero) grades must be submitted in writing to the instructor within a week after the grades are announced. After that period the grade stands.
3. Please note that a student will not automatically receive an “I” grade when he/she misses some work, or misses the final exam. An “I” is given to those students who have a passing average at the time of the ‘incident’. I grades must go through a formal approval process and must be based on extenuating or emergency circumstances according to UNCC policy.
4. Submission of work: It is the student’s responsibility to ensure that the instructor has received work submitted. This is especially important when work is submitted electronically.
   a. If you use email, insure that you keep a copy of the sent email, and ask for a ‘read receipt’.
b. If submitting via our online course site Moodle, always keep a copy of your work in your Moodle.

5. Communication Protocol:
   (a) Questions, Comments, and Requests
   - For any questions or clarification of class material, please ask them on Canvas (I need help discussion board) or via email whenever possible. Everyone in the class is encouraged to help answer the questions. If satisfactory answers do not emerge, the instructor will answer.
   - For any comments or requests, please send email to the Instructor and TA.

   (b) Canvas
   - Announcements will be posted in Canvas. Make sure to check the assignment area frequently enough to stay informed.

   (c) Emails
   - Each student is given an email account by UNC-Charlotte. This is the account that will be used by your instructor. Changes to class assignments or other course information will be posted online and may sent to you. Check your email daily. Do not send email to your instructor from any other account, as it will be considered spam, and be deleted.
   - Please use Canvas, not emails, for general questions, unless you wish to keep your questions or comments private.
   - When emailing your instructor, please use a specific subject line starting with "ITCS 6155: ", e.g., "ITCS 6155: Exercise 1".
   - The instructor will reply to legitimate email inquiries from students within 48 hours with the exception of weekends or university holidays. If you do not receive a reply within this period, please resubmit your question(s) or call your instructor. Leave a message if necessary. Cell: 980-234-0042.
Student Responsibilities:
Please refer to University Policy 406 - The Code of Student Responsibility, http://legal.uncc.edu/policies/up-406, for specific information. In addition to the responsibilities specified by the University, for this course, it remains the student’s responsibility to be aware of enrollment status, assignment due dates, changes to the syllabus, and deadlines for the UNCC academic calendar. Each student is responsible for his/her attendance and properly withdrawing from the course if necessary.

Disclaimer
This syllabus is intended to give the student guidance in what may be covered during the semester and will be followed as closely as possible. However, the professor reserves the right to modify, supplement and make changes as needed.

Good luck in class! I am looking forward to working with you this semester, and sharing my knowledge.

After reading this syllabus, please let me or our TA know if you have any questions – please contact me by email or phone with questions.