SPECIALIZED CERTIFICATE IN DIGITAL HEALTH:
Data Science for Digital Health Syllabus

Course 2: Data Science for Digital Health

Course Number: GLBH-40021
Section ID: 148968
Units: 3.00
Start Date: 06/22/2020
End Date: 09/12/2020

Instructor Information

Name: Hobson Lane, MS
Questions: Pose all questions on the discussion forum: ucsdextension.instructure.com/courses/805
Communication Policy: Instructor will respond to emails at dhlane@ucsd.edu within 24 hrs

Welcome

Welcome to the Data Science Digital Health online course. I hope you are as excited as I am to explore all the ways that Data Science can be used to advance health care, improve outcomes, and even save lives! Don't worry if you've never programmed using python, used statistics to diagnose disease, or trained a machine to prioritize public health interventions. We'll show you how to take advantage of all the amazing tools of Data Science to do all this and much more. You will start by using a spreadsheet program you are already familiar with, such as Libre Office or Excel. Later you will learn how to automate your work using the Python programming language. By the end of the course you will be implementing state of the art Data Science and Machine Learning models.

Course Description and Goals

This is the second course of the Digital Health Specialized Certificate. The course has been crafted by experts with deep experience in applying data science to healthcare. And because this field is evolving so rapidly, we have developed a chatbot that you can use to contribute to the course material. By asking the chatbot questions, rating its replies, and suggesting improvements to the answers, you will help us keep the course up to date.

Course Purpose and Prerequisites

Intended for those with a background or interest in healthcare transformation, data science, eHealth, public health, IT administration, engineering and regulatory affairs.

Course Objectives

Upon completion of this course, the student will:

- Understand the impact of Digital Health on the healthcare ecosystem
- Apply critical thinking to Data Science problems in the Digital Health industry
- Assess opportunities for Healthcare innovation and disruption with Data Science
- Train predictive machine learning and Data Science models on common healthcare problems
- Assess and explain the predicitons of Data Science and machine learning models

Course Materials/Textbooks

Required: *Healthcare Analytics Made Simple*
by Vikas Kumar, ISBN: 9781787286702
Recommended: *Natural Language Processing in Action*
by Hobson Lane, Cole Howard, Hannes Hapke, ISBN: 9781617294631
Recommended: *The Book of Why*
by Judea Pearl, ISBN: 9780141982410

Course Overview

This course has 10 sessions. The topics are as follows:

Session 1: **Lecture 1: Data Science in Healthcare**

Session 2: **Lecture 2: Spreadsheet Data Science**

Session 3: **Lecture 3: Statistics & Privacy**

Session 4: **Lecture 4: Clinical Data Science & Machine Learning**

Session 5: **Lecture 5: Deep Learning & Computer Vision**

Session 6: **Lecture 6: Natural Language Processing**

Session 7: **Lecture 7: Bioinformatics & Genomics**

Session 8: **Lecture 8: Assistive Technology, Ethics**

Session 9: **Lecture 9: Healthcare Systems Modeling**

Session 10: **Lecture 10: Public Health & Epidemiology**

Online Course Structure

The course is organized using the course menu:

- **Announcements:** This is the first page you see upon entering your course. Your instructor will post weekly announcements and reminders here.
- **Introduction:** Contains an introduction to the course and instructor biography.
- **Syllabus:** Contains the course outline, learning objectives, weekly assignments and course details.
- **Lessons:** This section will have the instructor's weekly audio/image lectures. The lectures are self-paced and can be replayed anytime like a video or movie.
- **Discussion Board:** Questions pertaining to each lesson are posted weekly for you and your classmates to discuss and answer.
- **Office Hours:** The instructor will host weekly group discussions on Zoom.
- **Assignments:** Exercises will include python programming assignments in Jupyter notebooks (online resources provided) for data manipulation, visualization, and predictive analytics). Quizzes and a final exam will allow you to measure your progress and reinforce your learning. You will also be able to provide feedback to the instructor and university in a course evaluation.
- **Resources:** Additional readings and handouts, website links, and PowerPoint presentations.

Technical and Policy Memo

Week	Assignments	Grades
0	**View lecture 0: Introduction to Course and Certificate** Assignment #1: Discussion Board	
1	**View Lecture 1: Data Science in Healthcare** **Assignment 1:** **Discussion Board**: Describe the similarities and differences between Artificial Intelligence, Machine Learning, and Analytics. Use each term in a sentence. **RR**: *Healthcare Analytics Made Simple*, Chapter 1 "Introduction to Healthcare Analytics" **SR:** *Natural Language Processing in Action*, Chapter 1 "Packets of Thought (NLP Overview)" **SR:** Jupyter Notebook Tutorial by Karlijn Willems: **bit.ly/ucsd-jup**	**Discussion**
2	**View Lecture 2: Spreadsheet Data Science** **Assignment 2:** **Discussion Board**: Give one example of a machine learning model type and the kind of problem it might be used for (hint: p. 53, 54 and 57 describe 3 model types). **RR:** *Healthcare Analytics Made Simple*, Chapter 3 "Machine Learning Foundations" **SR:** *Natural Language Processing in Action*, Appendix "Intro to Machine Learning" **SR:** Simple Linear Regression by Jason Brownlee: bit.ly/ucsd-sheet	**Discussion** Quiz Mini-Project
3	**Vie**w Lecture 3: Statistics **&** Privacy **Assignment 3:** Discussion Board and Quiz **Reading:** 1. **Required**: *Introduction to Healthcare Analytics*, Chapter 3 "Machine Learning Foundations" 2. **Required**: Veritasium YouTube Explanation of Bayes Rule using Common Sense: bit.ly/ucsd-bayes 3. Suggested: Natural Language Processing in Action, Chapter 2 "Build Your Vocabulary (tokenization)" 4. Suggested: *Statistical Thinking in Python*, DataCamp tutorial: bit.ly/ucsd-eda	Discussion Quiz

4	**View Lecture 4: Clinical Data Science & Machine Learning** Post to **Discussion Board**: Share the challenges you had in building a machine learning model to predict diabetes prediction models. Can you think of any ways to improve its accuracy? **Reading:** 1. **Required**: *Introduction to Healthcare Analytics*, Chapter 7: "Making Predictive Models in Healthcare" 2. Suggested: "AI outperforms doctors diagnosing breast cancer,"" BBC News: bit.ly/ucsd-bc 3. Suggested: Natural Language Processing in Action, Chapter 3 "Math with Words (TF-IDF vectors)	Discussion Quiz Mini-Project
5	**View Lecture 5: Deep Learning & Computer Vision** Post to **Discussion Board**: Which features were most helpful when creating a deep learning model on the deep learning playground? How many total neurons and layers did you need to achieve high accuracy on the spiral classification problem? If you used more features were you able to reduce the number of neurons required to fit to the spiral data well? **Reading:** 1. **Required**:*Introduction to Healthcare Analytics*, Chapter 9: "The Future – Healthcare and Emerging Technologies" 2. Suggested: Natural Language Processing in Action, Lane et al – Chapter 5 "Baby Steps with Neural Networks" 3. Suggested: "Artificial intelligence in healthcare: past, present and future" Jiang et al: (bit.ly/ucsd-ai-survey)[https://bit.ly/ucsd-ai-survey] 4. Suggested: View 8 minutes of Thinking Tools, 56:44 to 1:04:00 by Daniel Dennet	Discussion Quiz

6	View **Lecture 6**: Natural Language Processing Post to **ucsd.qary.me**: Ask TA Bot 5 questions about Data Science, Machine Learning, Python, or Digital Health: ucsd.qary.me. Try to think of questions that a healthcare practitioner would have and questions that cannot be answered using a search engine like Duck Duck Go. **Quiz**: Describe some unstructured datasets you've seen that Natural Language Processing could help with. **Reading**: 1. **Required**: NLP & Healthcare: Understanding the Language of Medicine 2. Suggested: Natural Language Processing in Action by Lane et al – Chapter 7 "Getting words in order with convolutional neural networks (CNNs)" 3. Suggested: Natural Language Processing in Action by Lane et al – Chapter 8 "Loopy (recurrent) neural networks (RNNs)"	Discussion Quiz Mini-Project
7	**View Lecture 7: Bioinformatics & Genomics** **RR: "Biopython tutorial"** **RR: "Contribution of GWAS..."** by Mansiaux et al, 2012, bit.ly/ucsd-gwas **SR:** Intuition Pumps and Other Tools for Thinking by Daniel Dennet p. 185-217 (pumps 33-37 Tools about Evolution)	Discussion Quiz
8	**View Lecture 8: Assistive Technology, Ethics** **RR:** *The Book of Why: The New Science of Cause and Effect* – Chapter 7 "Beyond Adjustment" **RR:** **SR:**	Discussion Quiz Mini-Project
9	**View Lecture 9: Healthcare Systems Modeling** **Exercise 9:** Excess readmission rate modeling **RR:** *Data Science for Healthcare: Methodologies and Applications* – p. 347-365 "Leveraging Financial Analytics.." **SR: "**Modeling Organizational Deerminant of Hospital Mortality", Al-Haider and Wan 1991, bit.ly/ucsd-hosp1991 **SR:** *Biomedical Informatics,* Shortliffe et al – Ch 12 "EHR"	Discussion Quiz
10	**View Lecture 10: Public Health & Epidemiology** **RR:** *The Book of Why: The New Science of Cause and Effect* – Chapter 5 "The Smoke-Filled Debate" **RR:** **SR:** *Biomedical Informatics,* Shortliffe – Ch 15 "Pub Health"	Final Exam