Data Science Picodegree

HURU School

10/15/2020

Overview

HURU School's Picodegree in Data Science will provide you a good foundation to become a world class data scientist. Through dynamic student led lectures, practical work and group discussions and presentations, you will gain exposure to the latest methods, techniques, and tools. The course will equip you the student with concepts and data science principles with a focus on today's cutting edge tools.

Course Objectives

• To impart an in-depth understanding of the interplay between data and data science tools. This includes (i) emphasizing the ways in which data insights can "scale up" to affect your organization, as well as (ii) understanding how data science has evolved to shape the business and scientific landscapes.

• To think creatively about major questions in data science and refine students' skills in posing ethical approaches to practical data science application.

• To appreciate the role of data collection, curation and storage for further data analysis and insight development.

• To achieve a working knowledge and level of comfort with cutting edge data science tools R and Python for science and business.

Instructors

Main Instructor:

Name: Joshua Oluoch

Email: datascience@huruschool.org

Office Hours: Friday 9:00am - 10:00am

Course Duration:

• 4 Months

Meeting Times:

9:00-10:30 AM, Mondays, Wednesdays - Regular Class

7:00- 8:30 PM, Mondays, Wednesdays - Evening Class

Requirements:

- Bachelors Degree / Equivalent work experience
- Statement of interest (Outlining familiarity with data science and why you want to enroll in the course)
- Laptop/Desktop Computer - with enough processing power to run data analysis locally

Learning Outcomes

Throughout the semester, students will learn to:
- Think about the ethical implication of the skills they will acquire.
- Envision and execute impactful data science pipelines.
- Prepare professional data science presentations.

Text:

There will be no course textbook; instead we will rely heavily on local case studies based on students data and resources from:
- R Programming - https://tidyr.tidyverse.org/
- Python Programming - https://pandas.pydata.org/

Grading:

Students will be evaluated based on attendance, presentation and analysis of the class project ("info update" presentations (10%), group/individual proposals: 10%, group/individual final presentation: 20%, individual online reports:30%, Attendance: 30%)

Class project:

You will work individually or in groups to design and propose a data science project. Students will be required to present this project, inculcate instructor feedback,carry out the project to completion and turn in a 2 pages single-spaced (max) writeup for their final project.

Course outline:

Week 1: Introduction to Data Science, Instructors, Students and course expectations

Week 2: Introductions to R and Python and their respective interfaces

Week 3: Introduction to Data Ethics and Numerical forms of data (Qualitative vs Quantitative)

Week 4: Data Collection (Primary vs Secondary)

Week 5: Group Info-updates based on student data interests

Week 6: Descriptive and Manipulation Data Analysis Tools - Using R and Python Libraries (Dplyr and Pandas etc)

Week 7: Scientific Data Analysis, Data Visualization and Statistics in R and Python

Week 8: Introduction to Big Data and Databases (SQL and Non SQL), Distributed systems such as Hadoop, Pig, Spark etc.

Week 9: Introduction to Machine Learning: Predictions, Supervised and Unsupervised Learning.

Week 10: Intro to Deep learning.

Week 11: Student Data Science Project Details and Brainstorming.

Week 12: Proposal Defense .

Week 13: Project Implementation.

Week 14: Final Report Writing.

Week 15: Final Presentations.

Week 16: Graduation!

Welcome to a unique learning experience at HURU!