OptumLabs Data Science Internship Program  
University of California, San Francisco  
Bakar Computational Health Sciences Institute  
Summer 2020

Program summary

This program will appeal to those student/trainees who wish to apply their Data Science training into the field of biomedicine and health care, an area ripe for innovation.

The UCSF Bakar Computational Health Sciences Institute (BCHSI) administers the OptumLabs Data Science Internship Program (“Program”). This Program aligns with UCSF’s core Research Computing Capability initiative at UCSF known generally as the Information Commons. The Program provides opportunity to develop technical solutions to enable researchers leverage the latest computing technologies in analyzing clinical data. The data sets currently in development for Information Commons include structured Electronic Health Records, unstructured clinical notes and images. There are several types of projects underway for Information Commons:

- Addition and de-identification of new data types from new sources (notes, images)
- Extraction of concepts from unstructured clinical text
- Development of effective information visualization methods/tools
- Development of user friendly tools to access, analyze and correlate data

We are currently in the third year of a five-year program to build this infrastructure and extract Information from several data sources in support of our mission.

Intern Opportunities

We seek 1-2 students from across the University of California campuses to participate in the Program. During their summer internship period (extendable part-time into the Fall by mutual agreement) the students will carry out specific projects that build upon efforts to create the infrastructure that supports Information Commons. Specific tasks may include:

- Development and automation of pipelines to extract data in the appropriate format for analysis of multi-factor data sets.
- Development of Natural Language Processing (NLP) tools to retrieve information from free-form text in the clinical notes; as well as performing and optimizing that information retrieval
- Development or configuration of information visualization tools, such as visualization of concepts extracted from clinical notes or visualization of patient timeline based on a variety of data sources.
• Development or configuration of information annotation tools, to enable data curation by medical content experts.

Desired Background

• Substantive coursework in Data Science and Computer Science, specifically related to Information Retrieval at scale
• Coursework in or demonstrated interest in human biology or pre-med fields
• Undergraduate rising Juniors and Seniors, and graduate students, will be given preference

Internship Details

• Interns are expected to work full-time (40hrs/wk) for a period of 3-4 months over June – September (specific schedule and dates flexible)
• Interns will be based at University of California, San Francisco, Mission Bay Campus
• Some work may be done remotely, but interns are expected to be physically present for meetings throughout the internship period
• Housing will not be provided
• Interns must be eligible to work in the US during the internship period
• Interns will be paid at a rate of $29/hr plus benefits

To Apply

Please send your CV along with a cover letter to the Program Contact, and include “OptumLabs Internship” in the subject line:

Angela Rizk-Jackson, PhD
Director of Operations
UCSF Bakar Computational Health Sciences Institute
Angela.Rizk-Jackson@ucsf.edu

In order to ensure consideration for the Program, Applications are DUE by APRIL 1, 2020