ERIC SUN

@ ejs336@cornell.edu 720-261-8304 suneric98.github.io github.com/suneric98 linkedin.com/in/suneric98

EXPERIENCE

Ankura

Machine Learning / Backend Intern

June 2019 – Aug 2019 Washington D.C.

- Used **python (sklearn, nltk)** to build a Random Forest text classifier with 96% accuracy and 93% recall
- Used **python (re, multiprocessing)** to cut down 6 hours of sentiment analysis to 2.5 hours
- Implemented an application that collected emojis from documents and performed sentiment analysis **(python)**
- Developed an email header parser that detected the language of an email **(Java)**

Guidepoint

Data Analyst Intern

June 2018 - Aug 2018 New York, New York

- Made a web scraper (**python**) to gather medical sales data
- Cleaned and mapped data using **MySQL** and **Excel** to project healthcare company revenues.
- Used **R**, **python**, and **Excel** to analyze data for significant points. Created graphs to indicate trends and developed a prediction methodology with 2% error.

Columbia University

Instructor Assistant for Secondary School Field Research Program

July 2017 - Aug 2017 Lamont-Doherty Research Center

- Instructed 7 high school/college students in assisting research with Dr. Elizabeth Corbett and Dr. Dorothy Peteet
- Analyzed the carbon density of Piermont marsh using geological cores, **R** and **Excel**
- Instructed students in **Matlab** and **Excel** to create graphs and an interpolated depth map of the marsh

SKILLS

Programming

Experienced:

Python • Java • R • MySQL • Javascript • OCaml

Familiar:

LaTeX • Matlab • C • C++ • Swift • React

Other Skills

Word • Powerpoint • Excel • Outlook

Intermediate Chinese

EDUCATION

Cornell University

BA in Computer Science, Statistics

Cum. GPA: 3.412

Aug 2016 – May 2020

Tappan Zee High School

Aug 2012 – June 2016

PROJECTS

Cornell Data Science Project Team

Fake News Detection

- A project that classifies the relevance and stance of an article (**python**)
- Developed visualizations of our models (**Javascript D3**)

Wikipedia

- A project that classified the hierarchy of related Wikipedia articles **(python)**
- Created a force graph visualization displaying a random walk through related Wikipedia articles **(Javascript D3)**

Personal and School Projects

McDonald's Country Map (Personal)

- **R** program that scrapes the list of countries McDonald's has been in and animates them on a map from 1940 – Present

Texas Hold 'Em (School)

- **(OCaml)** Developed the UI for a fully functioning poker game with support for AI and multiplayer. Played in terminal with simple graphics

COURSEWORK

Teaching Assistant

STSCI 2150: Introductory Statistics

Relevant Coursework

CS 2110: Obj Oriented Programming
CS 2800: Discrete Structures
CS 3110: Functional Programming
CS 3410: Systems Programming
CS 4410: Operating Systems
CS 4780: Machine Learning
CS 4820: Intro Analysis to Algorithms
STSCI 4030: Linear Models w/ Matrices