Soon Jae Kwon

Philadelphia, PA 19104
Contact: (415) 359 6824 | kwonso@seas.upenn.edu

soonjae.works
github.com/kwonsjay
linkedin.com/in/soonjae

Education

- **University of Pennsylvania, School of Engineering and Applied Science (SEAS)** — **Philadelphia, PA**
  Master of Computer and Information Technology (3.86/4.00) — May 2020
  Master of Biotechnology: Biopharmaceuticals/Engineering Biotechnology Track — May 2016
  Bachelor of Science in Engineering (BSE) in Bioengineering — May 2013
- **Relevant Coursework**
  Algorithms and Computation, Data Structures and Software Design, Programming Languages and Techniques, Algorithms for Big Data, Big Data Analytics, Computer Systems

Technical Skills

Proficient: JavaScript, Python, R
Intermediate: Bash, C, Java, Ruby/Rails, SQL

Languages

Fluent: English, Korean
Conversational: Japanese

Work Experience

- **Bioinformatics Engineer** — **July 2018 - September 2019**
  Corestem — *Seoul, South Korea*
  - Developed custom pipelines for timeseries and stratification analyses of genomic data (R, bash, JS)
  - Trained machine learning models based on shrunken centroids for feature extraction (R)
  - Guided Phase III clinical trial design for statistical power in hypothesis testing
- **Full Stack Developer** — **June 2016 - July 2018**
  ChunLab — *Seoul, South Korea*
  - Developed a full stack bacterial DNA identification service, served as an API (LAMJ, AWS)
  - Maintained public microbiome portal EzBioCloud (refactoring, request optimization, deployment)
  - Developed a data tagging system and interactive genomic data visualizations (D3)
  - Implemented workload scaling using AWS EC2 and Sun Grid Engine
- **Graduate Intern** — **June 2015 - August 2015**
  Genentech — *South San Francisco, CA*
  - Assisted in building a knowledge management program by scraping regulatory databases.
  - Built web presence of the PTR Biologics Program Management Office on Roche's intranet (Node.js)
  - Developed automation tools for team members (Python, Google Apps Script)

Projects

- **Amazon Review Star Rating Prediction using Neural Networks** — **Course Project**
  Classification of Reviews using Context-less Bag of Words — *November 2019*
  - Performed feature reduction to identify words explaining most of the observed variation (Python)
  - Created a baseline random forest classifier using reduced featureset, review times, and review length
  - Built and evaluated the performance of a neural network against the baseline model (Keras, TensorFlow)
- **Serverless Web App: WhatTheGoon** — **Personal Project**
  A Guide to the Korean Mandatory Draft for Third-Culture Kids — *September 2019*
  - Developed a Single-Page Application (SPA) using React.js and react-router for the front-end
  - Implemented a serverless back-end using AWS API Gateway, DynamoDB, and Lambda
  - Offloaded user management and authentication by using AWS Cognito and IAM
- **Trend Identification using Legacy Databases** — **Course Project**
  Analysis of FDA Adverse Events Reporting System Databases for Significant Trends — *May 2016*
  - Processed high dimensional data to calculate correlation scores between features of interest (R)
  - Performed search for drug interactions with the most reported adverse events and complications
- **Single Page Application for Live Audience Polling: Troll Poll** — **Personal Project**
  Implementation of Poll Everywhere's Live Audience Polling Feature — *January 2014*
  - Developed a Single-Page Application (SPA) using Ruby on Rails back-end and Backbone.js front-end
  - Implemented Twilio and Pusher APIs for instant data reload upon receiving votes via text