

EDUCATION

Indiana University, Bloomington, IN
Master of Science in Data Science

Expected May 2018
Current GPA: 3.78 / 4.0

Handong Global University, Pohang, South Korea
Bachelor of Science in Computer Science & Management, Cum Laude
Merit Scholarship (Top 1% on Spring 2015)
Community Leadership Training Team Leader Scholarship

August 2016
Cumulative GPA: 3.94 / 4.5
Fall 2014 – Fall 2015
Fall 2014

AWARDS

First Place in 2017 Indiana Medicaid Data Challenge - Data Analysis

October 2017

- Declared winners of the Data Challenge hosted by Indiana Chapter of HIMSS and Regenstrief Institute
- Visualized potential 'under-served' areas in mental health care in the Indiana state via Tableau
- Constructed 94% accuracy classification model using Random Forest that detects 'under-served' areas

PROJECTS

Reinforcement Learning: Comparative Study of Non-Stationarity

January 2017 – May 2017

- Designed a simple domain "Bus Gridworld" to test non-stationary environment in the tabular setting which is undiscounted, episodic task with start and goal states composed of bus route and walking path
- Demonstrated Dyna-Q architectures adapt well to the non-stationary environment
- Discovered that under epsilon-greedy policy, Sarsa agent also adjusted well to the changing environment

Ambulance Siren Detection

January 2017 – May 2017

- Proposed a novel classification method that enables rapid classification
- Constructed 89% accuracy classification model that detects the ambulance siren in noisy traffic setting
- Reduced the dimension of the siren signal by selecting ten basis vectors found from Non-negative Matrix Factorization (NMF) on the one-second chunks of Short-Time Fourier Transformed (STFT) data

Single-Cell Classification

January 2017 – May 2017

- Designed 96.5% accuracy classification model using SVM in classifying cell types from RNA-Seq data obtained from brain cells of mice that carry the cognitive and sensory functions of the mammalian body
- Suggested a 99% reduction in storage by reducing the dimension using Principal Component Analysis (PCA)

Heart Disease Analysis

September 2015 – December 2015

- Developed 82% accuracy model that predicts the presence of heart disease using the heart disease database of 920 patients with 14 distinctive features
- Graphically compared the performance of SVM, Naïve Bayes, and Decision Tree classifiers using SPSS

WORK EXPERIENCE

Machine Learning Intern, Biology Lab at Handong Global University

March 2016 – June 2016

- Constructed 0.895 area under curve (AUC) Logistic Regression model in classifying the movement of protein into chloroplast or mitochondria using N-terminal 1-20 of transit peptide sequence data
- Applied genome sequencing program GLAM2 to find the common motif in the sequence

Teaching and Research Assistant, Handong Global University

March 2015 – June 2016

- Version-controlled group project repositories using Github, handling conflicts when merging working nodes
- Participated in Open edX project by configuring the full-stack in three ways: using Virtual Box and Vagrant, from scratch on a single Ubuntu 12.04 server, Amazon Web Services pre-installed image on a single server
- Organized structure of computer programs teaching materials using Jupyter Notebook
- Guided students with problem-solving and assignments for Data Structures and Java Programming

Python Textbook Translator and Instructor, Handong Global University

February 2015 – February 2016

- Published a textbook with Dr. Youngsup Kim, the former president of Korea OCW Consortium, translating "Introduction to Computation and Programming in Python" by John V. Guttag from English to Korean
- Initiated several Python camps for 40 students in three universities and lectured on Python programming

TECHNICAL SKILLS

Data Tools: Numpy, Pandas, Scikit Learn, Tableau, SPSS

Languages: Python, R, SQL, Java, C/C++, Bash, Matlab

Operating Systems: Red Hat, Ubuntu, Linux Mint, Windows

Version Control, Editing: Git/GitHub, Vim, Latex

Web Development: Java Server Pages(JSP), JavaScript, HTML, CSS