Ethan Kim

(617) 758-9553 | ethan_kim@college.harvard.edu | 308 Amory St, Jamaica Plain, MA 02130
Github: https://github.com/ethankim00/ | Website https://ethankim00.github.io./

EDUCATION

Harvard University   Cambridge, MA
Leave of Absence Fall 2020
A.B Candidate in Statistics (Data Science Track), Secondary in Computer Science GPA: *3.76 /4.00.*   *Expected Dec '21*
Coursework: Advanced Topics in Data Science, Algorithms and Data Structures, Data Science, Time Series, Linear Models, Probability, Inference, Introduction to Computer Science, Multivariable Calculus, Linear Algebra,

Boston Latin School   Boston, MA
GPA: *4.52*, SAT: 2370/2400 (V800, M800, W770)   *May '17*
Activities/Awards: Varsity Soccer League All Star, National Merit Scholar Finalist, National Honor Society

SKILLS

Technical Skills: Python, Scikitlearn, TensorFlow, Matplotlib, Seaborn, Dask, SQL, R, AngularJS, D3.JS, git

RELEVANT EXPERIENCE

BGI LLC   Boston, MA
  Data Science Intern   *Sep 20 - Present*
- Use Natural Language Processing to analyze cockpit recordings and improve aviation safety

BrainSpec   Boston, MA
  Data Science Intern   *May 20 - Aug 20*
- Developed visualizations to present Magnetic Resonance Spectroscopy data from brain scans for clinical use
- Created Python data processing pipeline for analysis of spectroscopic data

Harvard College Data Analytics Group   Cambridge, MA
  Engagement Coordinator   *Aug '20 - Present*
- Manage and provide technical support to multiple Data Science consulting case teams
- Work with other members of the board to source clients for fall 2020

  Data Science Consulting Team Leader   *Jan '20 – May '20*
- Led team of six on Data Science research project for a sports analytics company
- Explored low latency predictive models for sports matches
- Facilitated launch of new club focusing on helping companies and non-profits derive value from Data
- Interviewed potential members and developed technical standards for the team

  Coronavirus Response Team
- Created visualizations for a high impact COVID paper (JAMA Internal Medicine)
- Performed initial data processing and consulted on effective visualization techniques

Harvard Open Data Project   Cambridge, MA
- Completed curriculum on data scraping, statistical analysis and data presentation.   *Feb '19 - Present*
- Implemented tools to visualize trends in menu data scraped from the Harvard Dining Service website

PROJECTS

Impact of Weather on Social Distancing
- Scraped county level weather data to evaluate relationship with social distancing indices
- Explored effect of weather on Coronavirus reproductive rates

Predicting Types of Crime in Boston
- Using spatiotemporal data to predict category of crimes from location
- Achieved significant increase in accuracy using census and location data and Random Forest models

Solving a Scrabble Like Word Game
- Designed data structures to efficiently predict playable words for a scrabble like word game
- Created training tool to help improve at game

Detection of Cell Cycle Genes and Regulators in S. Cerevisiae
- Time series analysis on mRNA expression data in yeast to detect cyclic genes
- Correctly identified key cell cycle proteins and regulators

Training an AI to play Incan Gold
- Implemented playable version of the board game, Incan Gold, using Python, Flask and JavaScript
- AI trained on winning moves from thousands of games achieved 150% increase in win percentage

LEADERSHIP AND ACTIVITIES

Harvard Academic Resource Center   Cambridge, MA
Peer Tutor   *Sept '18 - Present*
- Tutored Students in Math 21A Multivariable Calculus and Stat 111, Introduction to Statistical Inference

Interests: Sports Analytics, Reading, Football, Basketball, Harvard Club Soccer