

Apoorva Havanur

ahavanur@andrew.cmu.edu | ahavanur.github.io | 408-315-7249 | U.S Citizen

Education

Carnegie Mellon University, 2014 - 2018

B.S in Statistics and Machine Learning, additional major in Economics, minor in Computer Science
Quantitative Science Scholars Program (QSSS), Alpha Kappa Psi Co-ed Professional Business Fraternity
GPA: 3.71/4.0

Relevant Coursework:

10-601 Introduction to Machine Learning, 11-441 Natural Language Processing, 36-462 Data Mining, 36-402 Advanced Methods in Data Analysis, 15-210 Data Structures and Algorithms, 15-150 Functional Programming, 73-374 Econometrics II

Work Experience

Facebook

Data Science Intern, Growth Contacts Team

May 2017 – present

- Developing new identity verification techniques

IBM Watson Health

Data Science Intern, Innovations Team

May 2016 – Aug 2016

- Implemented and deployed risk model for coronary heart disease for use by healthcare providers across the country using Java and MapReduce, using records from more than 55 million unique patients stored in Hbase
- Designed novel fuzzy string matching algorithm for use in medical record standardization platform that increased detection of mismatches by 600%, while decreasing the number of false positives by over 50% compared to previous algorithm.
- Added procedure revenue metrics in internally generated data quality reports to analyze revenue distributions within patient records
- Received commendation for clarity and impact of work from Watson executives after presenting findings during the summer finale event in Austin, Texas.

Carnegie Mellon University

Research Assistant

Aug 2015 – Jan 2016

- Worked with Professor Joachim Groeger to develop model for taxi drivers to optimize route and rider selection using data collected on more than one million taxi trips.
- Used R to cleanse data and create distributions of fare amounts, pickup/dropoff locations, and duration
- Developed relational database in Postgres from original csv files to allow for SQL querying

Projects

Pittsburgh Civic Light Opera Donor Report

- Created interactive data visualization tool to aid non-profit theater company in analyzing more than 20 years of donation and ticketing data using R Shiny and ggplot (<https://ahavanur.shinyapps.io/clo-research-demo/>)
- Created prediction models for suggested donation amount using ticketing and past donor information, and presented findings at undergraduate research colloquium (Meeting of the Minds).

Tartan Data Science Cup – 1st Place Overall

- Analyzed Citibike dataset for CMU data science competition, working in team of 3 against more than 100 other students.
- Implemented Naïve Bayes classifier to determine gender of unknown riders
- Modeled bike traffic using graph theory and network tools in Python and determined most popular locations based on clustering coefficients and betweenness centrality of nodes.

MediMinder – Watson Health Hackathon

- Created interactive text-message based medication scheduling and reminder application for use by parents and children with chronic illnesses using Ruby on Rails and Twilio API.
- Utilized Watson Natural Language Classifier API to classify user inquiries about their medication and respond with appropriate information after querying database.
- Participated in a team of 4 against other undergraduate to PhD student teams across the country, coming in 4th.

Skills and Leadership

Resident Advisor

Aug 2015 – present

- Previously responsible for freshman floor of 33 residents and currently responsible for 40 upper-class residents
- Received training in conflict resolution, identifying problematic behavior and first responder training.
- Organized and ran dorm-wide events such as group outings, socials, and discussion groups.

President, Moneythink CMU

Sep 2014 - present

- Taught high school students in low-income areas of Pittsburgh lessons on financial literacy and good financial planning.
- Co-developed website (www.moneythinkcmu.org) as an informational and recruitment tool.
- Led board of 10 members into expanding the chapter in size and reputation around campus

Programming: Java, Python, R, (all proficient), SQL, SML, MapReduce, Hadoop/Hbase, Ruby on Rails, C (intermediate)

Other: Microsoft Office, Public Speaking, Conflict Resolution, Basketball, Standup Comedy