Karan Sarkar
978.908.8841. karan@sarkarfamily.org
7 Country Rd, Westford MA 01886

SUMMARY

Talented technical professional with a solid mathematics foundation with hands-on experience including in-depth knowledge of C, C++, Java, R. Looking for a summer internship where I can leverage my background in computer programming, mathematics, data analysis and fundamental research. While at high school, I was selected to a highly competitive yearlong research program at MIT. My research paper was published in the "International Journal of Game Theory".

Languages: C, C++, Java, Python, R, MATLAB, SQL, Powershell
Libraries: STL, Eigen, numpy, pandas, nltk, keras, matplotlib, bnlearn, ggplot2,
Tools: Eclipse, RStudio, Visual Studio, SQL Developer, MATLAB**,** GitHub
Courses: Data Structures, Algorithms, Data Analytics, Natural Language Processing, Data Mining, Machine Learning Intro to AI, Numerical Computing, Advanced Calculus, Econometric Methods for Big Data

PROJECTS

Healthcare Financial Inc. (Data Scientist) **August 2018 – October 2019**

- **Project 1 –** Developed rule mining algorithm to find high disability clusters
  - Generated candidate rules use pairwise support and lift
  - Constructed custom rule linked prefix-tree data structure to allow fast evaluation of rules
  - Used C++ STL containers such as unordered-map to obtain high performance
  - **Use of data structures reduced run-time by 75% from 4 hours to 1 hour**
- **Project 2 –** Constructed intuitive conversion from diagnosis descriptions to social security determinants
  - Used PubMed Word Embeddings to convert diagnosis descriptions to quantitative domain
  - Generated intuitive categories from spherical K-means
  - **Resulted in 0.25 Purity**
- **Project 3 –** Developed neural network classifier for disabilities
  - Pulled data across tables from MySQL server
  - Reduced overfitting through use of regularization and cross-validation.
  - Reduced feature set through Chi-Squared hypothesis testing and code truncation
  - **Reduced run time from 1 day to 1 hour**
  - **Improved AUC to 96% from 93%**

Rensselaer Polytechnic Institute, Troy NY (Student) **August 2016 – Present**

- Created C++ STL Style Graph Library
  - Used C++ template arguments to allow choice in internal representation i.e. vector vs hash-table vs linked-list
  - Created de-linked graph algorithm templates that allow dependency injection. Can be used to yield Djikstra's vs DFS vs BFS with just change of parameter
- Implemented Paxos Algorithm using Python TCP Programming
  - Maintains distributed log across many machines
  - Used appropriate data structures such as hash-table to maintain efficiency
  - Made use of Python threading library to maintain safety
  - Constructed flexible Event class to maintain extensibility

Partners Healthcare (Summer Software Intern) **July - August 2017**

- Created a Citrix Messaging Application that sent messages to a number of selected delivery groups using Powershell

KEY MILESTONES

Interviewed on national television on FOX news **2011**
Got a 800 in Mathematics for the SAT Math and 700 in SAT Verbal **at the age of twelve**

Program for Research in Mathematics, Engineering, and Science, Massachusetts Institute of Technology **2015-2016**
Wrote a research paper that was **published at a MIT conference and published in the International Journal of Game Theory.** The paper is called "On P-Positions in Modular Extensions to Nim" http://link.springer.com/article/10.1007/s00182-016-0545-7