Saumya Parikh

480-363-8084   saumya1510@gmail.com

EDUCATION

Arizona State University, Tempe, AZ   May, 2018
Master's in Computer Science, GPA: 3.22/4
Nirma University, Ahmedabad, India   May, 2016
Bachelor's in Computer Engineering, GPA: 7.12/10
Relevant Courses: Foundations of Algorithms, Adaptive Web, Data Visualization, Data Mining, Distributed Database Systems.

SKILLS

Proficient: RESTful API Development, Algorithms and Data Structures, Machine Learning, Data Visualization (D3.js, Tableau), Flask Microservices.
Familiar with: Mongo, Hadoop, ElasticSearch, Java Spring, AngularJS, AWS, Azure.
Programming Languages: Python, C, C++, JavaScript, SQL, HTML/CSS.

WORK EXPERIENCE

Bio-Rad Laboratories, Inc.   **Summer Intern - Data Science Team**
Hercules, CA   June 2017 - August 2017

- Worked cross-functionally in an agile environment with the Life Science team to assess return on investments of different projects and campaigns.
- Developed an automated pipeline to extract, integrate and present data from multiple sources to monitor daily campaign performances, and make strategic marketing decisions.
- Built a scraper to extract results from the knowledge panels of search engines like Google, Bing, DuckDuckGo and Yahoo to validate raw user input data that improved integration with legacy sources like BAAN by 45%.
- Designed and implemented an innovative solution to the problem of attributing cumulative revenues from SAP data to customers and campaigns in Salesforce tables and attained close approximations of campaign ROIs.
- Tech Stack: Hadoop HDFS, Impala, MS Azure, Power BI, Python, Pandas, Flask, MySQL, JavaScript.

Media IQ Digital Pvt Ltd   **Developer/Analyst Intern**
Bangalore, India   January 2016 - June 2016

- Collaborated with data scientists and developers to optimize ad campaign performances, and report insights to traders and clients.
- Built an automated batch data pipeline using Appnexus APIs to collect user information via tracking pixels and store it in S3 containers, providing analysts data for building custom audience segments and improve targeting.
- Compiled and analyzed A/B testing results to create custom audience segments based on demographics and geography, which increased the click-through and conversion rates for various campaigns.
- Improved campaign performances by 30-50% by performing exploratory data analysis, and implementing logistic regression and decision tree classification on existing performance data to get audience overlaps and market trends.
- Tech Stack: PostgreSQL, Python, Scikit-learn, Redshift, Hadoop, Java Spring, MS Excel, AWS S3.

PROJECTS

- **Visualizing Climate Change**: Illustrated the correlations between increasing carbon emissions, average temperatures, frequencies and intensities of meteorological disasters and decreasing ice extent in the context of climate change via various interactive online visualizations developed using D3.js and AngularJS.
- **Java Wikibooks Search Engine**: Built a web application with Flask and Jinja template, to query Stack Overflow content to an ElasticSearch server on AWS, and retrieve top matches from the scraped and parsed Java Wikibooks stored in a MongoDB collection.
- **QuizIt - Questions Recommendation System**: Designed an adaptive quizzing platform that records users preferences and performance to provide them suitable questions, and enables them to assess their progress via personalized visualizations. Developed the app with React and Node JS.
- **Geo-spatial Data Analysis**: In a team of 4, analyzed NYC Taxi data, using Apache Spark on a Hadoop multi-node cluster, to identify top 50 hot-spots within a given location and time-range by calculating the Getis-Ord statistic measure in Map Reduce paradigm.
- **Spotify Top Tracks Analysis**: Performed exploratory data analysis on various audio features of my most played tracks, collected via the Spotify API, and developed visualizations to illustrate the outliers, correlations between music genres and moods, and my short and long term listening trends.