
Professional Summary

- Proficient in Data Analysis/Analytics/Mining, Machine Learning, NLP techniques, Predictive Mining & Analytics.
- Worked in python and expertise in python libraries NumPy, SciPy, Pandas, Scikit-Learn, Matplotlib, Plotly, PySpark, Seaborn.
- Worked in R and expertise in R libraries dplyr, ggplot2, googleVis, tidyr, tm, R Markdown etc.
- Functional experience with concentration on Data Visualization (d3.js), Tableau (Creating Dashboards), Technical Training, Software Development methodologies, Reports, Letters.
- Oracle Certified SQL Associate and Profound knowledge of Database Design.

Specialization

Data Science, Big Data, Machine Learning, Natural Language Processing, Data Visualization, Statistics and Probability, Predictive Analytics, Data Mining.

Skill Set

Programming: C, SQL(Certified), Python(Proficient), MatLab, HTML, CSS, JavaScript (d3.js), JSON, Apache Spark.
Data Mining & Tools: R(Proficient), Tableau, Hadoop Eco System, Microsoft Office (Excel, Access, PowerPoint), LaTeX, NLTK.
OS: Linux, UNIX, Windows.
DBMS & Clouds: Oracle, MySQL, AWS, Microsoft Azure.

Portfolio

- GitHub: <https://github.com/abachu>
- LinkedIn: <https://www.linkedin.com/in/akashbachu/>

Relevant Experience

Research Experience:

- Twitter and Polls: Performed Sentimental Analysis on Indian General #Elections2014 and visualized data using ggplot2, GoogleVis, wordcloud using R and used R Markdown for reports.
- Taken a shot at Time Series U.S refugee and Global Terrorism Dataset to make visualizations and discover bits of knowledge on worldwide assaults utilizing JavaScript (d3.js) and Tableau.
- Worked on H-1B Visa petitions, created visualizations to derive business issues, made a working model and checked the accuracy of the model using classification techniques. Also, created a dashboard in Tableau.
- Strategies in handling large data sets quickly and efficiently by running MapReduce program in super computer and verify the efficiency of parallel computing and Hadoop.
- Created a Database application (Employment management service) using technologies: JSP, Servlets, HTML | Database: MySQL | IDE: Eclipse
- Worked on building a Movie Recommender system with an accuracy of 93% using Pandas and Apache Spark (Stand Alone Cluster), AWS and Azure.
- Working on building a Document Classifier using Natural Language processing and Deep Learning techniques on MIMIC-3 Dataset (master's project).
- Received Internship at MIT Sloan School of Management as a Data Scientist to work for a Research paper but offer letter rescinded due to time constraints.
- Next word prediction using unstructured data. It involves devising an algorithm that uses stored word phrases (n-grams) to predict the next word that will appear after a given phrase.

Professional Experience:

- **SQL Analyst in GlobalLogic Technologies Pvt Ltd (May 2016-Aug 2016):** Performed SQL operations like stored procedures, views, functions, complex joins, grouping, aggregation, nested subqueries and cursors. Also, Imported, exported and manipulated large data sets in multi-million-row databases under tight deadlines.
- **Working as a Data Analyst Intern at Town of Dartmouth (Sept 2017 – present).**

Activities:

- Volunteered Health Innovators meetup to assist them with python and help entrepreneurs and growing companies partner with investors.
- Secretary for a Non-Profit Organization (One Step Foundation). We help aspiring children to fulfil their goals.

Education

Master of Science, Data Science.
University of Massachusetts Dartmouth, MA, USA

Sept 2016