SAI AKASH

703-909-8564 | saiakash5@gmail.com | WA 98052 | linkedin.com/in/saiakash

PROFESSIONAL EXPERIENCE

Vedamsoft, Java Hadoop Developer   05/17- Present

- Currently working on data generated from Omnicell robotic arm XR2.
- Currently working on efficient way of exporting data from API to HDFS using Apache NiFi
- Installed and configured MapReduce, HIVE and the HDFS; implemented CDH3 Hadoop cluster on CentOS. Assisted with performance tuning and monitoring.
- Supported code/design analysis, strategy development and project planning.
- Developed multiple MapReduce jobs in Java for data cleaning and preprocessing.
- Created reports for the BI team using Sqoop to export data into HDFS and Hive.

University of Louisiana at Lafayette, Graduate Research Assistant   09/15 - 05/17

- Developed a novel Apache Flink application for processing a graph data stream that consisted of millions of interconnected nodes by collaborating with a PhD research fellow.
- Implemented sketching algorithms on graph with million nodes and maintained data integrity.
- Calculated various similarity measures such as Common Neighbours, Adamic-Adar, Preferential Attachment and Jaccard Coefficient using graph sketches resulting in less memory usage.
- Compared stream processing performance between Apache Flink, Apache Spark, and Apache Storm.
- Reduced application result time by a record 33% through data optimization using graph sketches.

University of Louisiana at Lafayette, Teaching Assistant   01/16 - 05/17

- Assisted undergraduate students in acquiring Python skills by conducting detailed lab sessions, resolving programming issues, and performed grading where I was exposed to various kinds of approaches to solve a particular problem.

Bharat Heavy Electricals Limited, Intern   05/15 - 07/15

- Developed a web portal named "Business Excellence Management Portal" with main objective is to evaluate task assignments and produced accurate analytical reports using Eclipse BIRT by creating an employee monitoring system and by coordinating with design teams and developers.
- Completed project within time constrain of 6 months which involved database development using complex SQL queries.

Mindtree, Intern   12/13 - 04/14

- Created and implemented schemas using SQL scripts and deployed on Microsoft Team Foundation Server, whose objective was to generate stored procedures along with 4 SSIS packages.
- Other responsibilities included updating and creating new reports using Microsoft SQL Server Reporting Services.
- Performed unit testing on all code and packages and improved workflow reach by 40%.

EDUCATION

Master of Science – May 2017
Computer Science
University of Louisiana at Lafayette
GPA-3.66 Focus on Algorithms, Data Structures, Web Development and Data Analytics

Bachelor of Technology – April 2015
Computer Science
Jawaharlal Nehru Technological University
GPA-3.5 *Focus on Programming, Data Structures and Algorithms.*

PERSONAL DEVELOPMENT

- Currently working on a project using Python Flask Application for maintaining multimedia files using AWS resources on Linux and utilizing Cloud9 IDE for development.
- Currently enrolled for Microsoft professional program focusing on Data Science and Machine Learning.

APPLICATION DEVELOPMENT PROJECTS

- **Information Retrieval System Project Using Map Reduce Algorithm:** Developed an application that processed 10K documents and retrieved ranked results for any user input query by applying Vector Space Model, utilizing Amazon web service EMR with Apache Hadoop libraries and increased execution time by 30% and reduced memory utilization by 40%.
- **Visualization of Fitbit Data Project:** Pinpointed various user locations using Google Maps API by introducing a Fitbit API equipped application and by visualizing data using Google charts and JavaScript.
- **Information Retrieval System Project Using Python:** Created an information retrieval system with extensive processing capability by exploring various data structures and other imported libraries in Python.
- **Twitter Stream Project:** Developed an application to fetch twitter data and by storing it in Hadoop Distributed File System (HDFS) based on user specific keywords.
- **Statistics of Movie Data Project:** Processed movie ratings datasets while producing various statistics including specific user ratings and genres by designing an application in Apache Spark using Scala and its libraries like Spark SQL for SQL queries and Brunel for visualization of data.

SKILLS

Languages: Java(Expert), JavaScript(Proficient), SQL(Proficient), Python(Intermediate), Scala(Intermediate), C++(Intermediate), and C(Intermediate).
Frameworks: Hadoop, Spark, Flink and Hive.
IDE: IntelliJ, Eclipse, NetBeans, and PyCharm.

ADDITIONAL

- Strong analytical skills (GRE 310/340).
- Quick learner
- String coding skills