Sabiha Hussain Barlaskar

8600 Brodie Lane, Apt 735, Austin, TX 78745
Phone: +1 (512) 975 8391
Email: sabihabarlaskar90@gmail.com

PROFESSIONAL EXPERIENCE

CAPGEMINI INDIA PVT. LTD   *November 2014 – December 2015*
Software Engineer

Integrated Data Warehouse Application   *August 2015 – December 2015*

- Ingested batch data from various data sources such as RDBMS using Sqoop into Hadoop Raw layer
- Checked the data for consistency and accuracy such as the count of data listed in the Primary key column, and assessing the quality of the data for validity, accuracy, completeness, consistency and uniformity using Hive queries
- Wherever the data was inconsistent, cleaned the data using Hive queries and Hive jars

  Technologies used: Hive, Sqoop, Hue, Java

Integration of Neo4j with Apache Spark to find strongly con nected components with Twitter data   *May 2015-July 2015*

- Imported Twitter data via Twitter API using access tokens and API keys into Neo4j using Python
- Cleaned the data for consistency
- Displayed all nodes strongly connected by the FOLLOWS relationship as a graph using Cypher query language

  Technologies used: Python, Cypher query language, Docker compose platform, Neo4j database

Detection of Fraud rings in e-commerce transactions   *February 2015 – March 2015*

- Loaded transaction data into Hive from flat files

- Identified fraudulent transactions by triangulating the location of the user performing the transaction through the purchase process with the IP address of the transaction, the mobile tower location of the user at the time of transaction, the address registered and the billing address of the user

- Executed Cypher query to filter out the transactions where IP address location and mobile location were different and displayed the fraud rings in Neo4j

  Technologies used: Hive, Hadoop pseudo node cluster, Java, Cypher query language, Neo4j database

TEXAS INSTRUMENTS   *January 2014-June 2014*

System admin Intern

- Developed Account Management Utility using LDAP for internal TI users which works as an aide for cleaning up of Unix accounts of TI employees
- Automated SAN Storage Migration using Perl

  Technologies used: Perl, Perl CGI, Unix commands, LDAP directory, JavaScript, HTML

SKILLS

Hadoop technologies like Hive, Pig, MapReduce, SQL technologies, Java, Python, Perl, MS Excel, Scala, R, Neo4j, Cypher query language, HTML, Dimensional modeling, PL/SQL, ER modeling, UNIX

EDUCATION

Udacity

Data Analyst Nanodegree   *March 2016- Present*

Gauhati University   Guwahati, India

Bachelor of Science in Computer Science   *June 2011*
79,7%
Was among the top 3 in a class of 80 students

Master of Computer Applications   *June 2014*
81.6%
Topped a class of 150 students

CERTIFICATIONS

- Achieved the DB2 certificate from IBM and earned the title "IBM DB2 Academic Associate".
- Oracle Database SQL Certified Expert

PUBLICATIONS

International Journal of Advanced Research in Computer Science, Volume 5, No. 7, September-October 2014 "**Creating Graph Databases From Relational Databases: An Implementation**"