

Contact

www.linkedin.com/in/mishra-shantanu (LinkedIn)
portfolio.com/ShantanuMishra (Portfolio)

Top Skills

MapR
Apache Kafka
Apache Spark

Certifications

Hadoop Fundamentals
Kafka Essential Training
Becoming a Web Developer: Full Stack vs Front End
Building Deep Learning Applications with Keras 2.0
Extending Hadoop for Data Science: Streaming, Spark, Storm, and Kafka

Honors-Awards

Technical Excellence Award
Certificate of Appreciation
First prize in Techniti
Quality Assurance Intern

Shantanu Mishra

Data Scientist at Zions Bancorporation
Fullerton, California

Summary

Highly motivated, focused, diligent, eager to learn new things, team player bringing more than 5+ Years of total IT Experience with excellent experience in design, analysis, development, testing, and deployment of enterprise, client-server, distributed, Bigdata, Internet/ Intranet based database applications.

Experience

Zions Bancorporation
Data Scientist
August 2018 - Present (1 year 8 months)
Greater Salt Lake City Area

Zions Bancorporation
Data Science/Data Engineering Intern
May 2017 - August 2017 (4 months)
Greater Salt Lake City Area

- o Developed and migrated BASH based data ingestion pipelines for Banking applications using Streamsets.
- o Developed generic python scripts to automate data ingestion stages of Streamsets application.
- o Developed unit test cases and designed workflow for mobile sign-on application's fraud detection module.

California State University, Fullerton
Data Center Analyst
January 2017 - May 2017 (5 months)
United States

- o Automated reports extraction from servers using Elastic Search
- o Developed dynamic reports for server logs and projected the health status in charts.

Amdocs
Big Data Developer

July 2015 - July 2016 (1 year 1 month)

Pune, Maharashtra, India

> Big Data Developer at AT&T

- Technologies: Hadoop, Mahout, Kafka, Spark, Hive, HBase, Pig, Sqoop, Hue, CDH, HDP, Linux (Red Hat and SUSE), Bash, Scala, Oracle, Teradata Data Warehouse Appliance, HP Vertica
- Developed and delivered BI reports in Big Data Snapshot team using Hadoop.
- Designed and developed ETL process for BI reports using Sqoop, Hive, Pig and Spark.
- Automated data extraction and transformation for CDR, DDR, DirecTV and SPI.
- Actively participated in design and implementation of big data backup and recovery.
- Automated Big Data applications' data ingestion and data reconciliation reports.
- Completed POC on Apache Spark execution on top of Amdocs Big Data Insights tool.
- Automated Data warehousing applications ingestion to Hadoop using Sqoop and bash scripts. Having Experience in writing HIVE queries & Pig scripts.
- Implemented high level designs in modules in an agile environment
- Developed modules and unit test cases in Junit for successful deployment on production.

> Big Data Product Development at Amdocs

- Technologies: Amdocs Big Data Insights, Hadoop, Spark, Hive, HBase, Pig, Sqoop, Hue, CDH, HDP, Linux (Red Hat and SUSE), Bash, Scala, Teradata Data Warehouse Appliance, HP Vertica
- Developed modules for enhancement of Amdocs Big Data Insights(ABDSI) tool.
- Designed and developed continuous integration of Big Data framework on Jenkins.
- Designed a process for indexing of data at source to improve framework build and deployment in production by 30%.
- Automated Big Data applications' data ingestion and data reconciliation reports.
- Actively participated and implemented Data Reconciliation process on top of Big Data.
- Experience in integrating HBase with Hive, Pig.

- Handle the JSON data using Hive, Pig and filter the data based on query factor.
- Worked on a POC to maintain Referential Integrity and foreign key concept on top of Hive using Amdocs Big Insights, Java, Hive, Pig, Oracle database.

Tata Consultancy Services

Data Warehouse ETL Engineer/ Big Data Developer

January 2012 - July 2015 (3 years 7 months)

Bengaluru Area, India

- Used Talend Data Studio for Extraction, Transformation, and Loading (ETL) of information from numerous sources like Flat files, XML documents, and Databases.
- Highly skilled in integrating Kafka with Spark streaming for high speed data processing.
- Worked with NoSQL databases like HBase, Cassandra and MongoDB for information extraction and place huge amount of data.
- Experienced in developing UDFs for Pig and Hive using Java.
- Hands on experience on developing UDF, Data Frames and SQL queries in Spark SQL.
- Exporting and importing data to and from Oracle using SQL developer for analysis.
- Experience in understanding the security requirements for Hadoop and integrate with Kerberos authentication and authorization infrastructure.
- Writing Hive join query to fetch info from multiple tables, writing multiple Map Reduce jobs to collect output from Hive.
- Developed the code for Importing and exporting data into HDFS and Hive using Sqoop.
- Used Kafka and Storm for real time data ingestion and processing.

e.Soft CONNECT

Project Trainee

May 2010 - August 2010 (4 months)

Lucknow Area, India

- > Designed and developed a web-based application for Smart City
- > Worked on JavaScript, Java, J2EE, MySQL, Apache Tomcat, and Netbeans to create an MVC architecture
- > Hosted web application using Tomcat and analyzed web-logs using Tableau and MySQL

Hindustan Aeronautics Limited

Quality Assurance Intern

June 2010 - July 2010 (2 months)

Lucknow Area, India

Worked as a quality assurance intern at HAL(Hindustan Aeronautics Limited) and designed modules for generating weekly and monthly performance reports.

Education

California State University-Fullerton

Master of Science (MS), Computer Science · (2016 - 2018)

Meerut Institute of Engineering and Technology

Bachelor's Degree, Computer Science and Engineering · (2007 - 2011)