

# Revanth M

## Big Data & DevOps Specialist

A passionate professional and a fast learner; having 5+ years of experience in building data intensive applications, tackling challenging architectural and scalability problems in Finance and Logistics. Currently helping DHL with Petabyte scale data pipelines and turning data into values and actions.



✉ revanthm93@gmail.com

☎ +919211192177

📍 Chennai, India

🌐 linkedin.com/in/revanthm93

📄 revanthm93

## SKILLS

Data Engineering Cloud Engineering AWS GCP Devops Datawarehousing ETL Machine Learning  
Data mining Exploratory Data Analysis Python Scripting Software engineering Data Visualization  
Real-time data Analytics Natural Language Processing Shell scripting Reporting hadoop administration

## TECHNICAL SKILLS

<b>Cloud</b>	AWS, Azure, Cloudera, MAPR, OpenShift	<b>BigData</b>	Hadoop, Spark, Map Reduce, Storm, Kafka, Sqoop, Hive, Elastic search, HBase, Pig, Logstash, Oozie, NiFi, Splunk
<b>Language</b>	Python, R, Scala, Java	<b>Vizualization</b>	Tableau, Kibana, Grafana, Qlikview
<b>Devops</b>	Docker, Kubernetes, Jenkins, Terraform, Ansible, Chef, Artifactory	<b>Machine Learning</b>	Pandas, Numpy, Scipy, scikit-learn, NTLK, Matplotlib, Tensorflow
<b>SQL</b>	Oracle, Teradata, MSSQL, MYSQL	<b>NOSQL</b>	MongoDB, HBase, Cassandra
<b>Scripting</b>	Shell, Python, Groovy	<b>Build</b>	Gradle, SBT, Maven
<b>Version Control</b>	Git, Bitbucket	<b>OS</b>	Linux, Windows

## WORK EXPERIENCE

### Big Data & DevOps Specialist Deutsche Post DHL

01/2019 – Present

Chennai, India

Projects Handled: Geocoding, Smart truck, DGF Finance

- Part of planning/migration team for Application & Datalake Migration from on-premise to cloud environment - AWS & Google cloud. Designed and implemented large scale data storage, data processing and data distributed systems.
- Provided operational and technical expertise to initiatives for the Big Data analytics platform, data and applications, addressing a broad range of technologies including Big Data / Hadoop, Spark, Kafka, HBase, Hive, Sqoop, Pig, BigQuery, BigTable, DataProc etc., on Linux hosted on Google or AWS cloud.
- Lead design of our data warehouse analytics cloud infrastructure, amalgamating data from multiple and disparate sources, and working with the various teams on implementation. Contribute in building an internal traffic tracking and analytics system for both real-time and comprehensive historical analysis with a self service visualization layer.
- Built a high performing real time pipeline solution for streaming data with 200X improved performance for compliance product using Kafka and Spark Streaming.
- Expertise in establishing DevOps across security (Kerberos), monitoring (OVO, Airflow, Oozie & others), extensively built CI/CD pipelines using Docker, Jenkins, ZooKeeper, Terraform, Ansible, Puppet, Chef. Proficient with Docker-Container orchestration.
- Performed EDA, applied machine learning NLP algorithms to various areas of recommendations, predictions, prescriptive analytics, optimizations, classifications, clustering and segmentation resulting in high accuracy of geocoding API & customer satisfaction.
- Well versed in Configuration and tuning production and development Hadoop environments with the various Hadoop components.
- Acting as a big data specialist, provided technical and process leadership for projects, defining and documenting information integrations between systems and aligning project goals with reference architecture.
- Managed, mentored and established a team of Junior Big data and devops

## **WORK EXPERIENCE**

### **Senior Big Data Engineer**

Comcast

10/2017 – 01/2019

Chennai, India

*Projects Handled: Xfinity Mobile, Broadband, Cables*

- Build the infrastructure required for optimal data extraction, transformation and loading of data from a wide variety of data sources using SQL, AWS big data tools. Process real-time streaming data using the ELK stack, Apache Storm, and Spark Streaming.
- Build analytics tools that utilize the data pipeline to provide actionable insights into customer acquisition, operational efficiency and other key business performance metrics.
- Developed 'go-configure' ETL frameworks with the help of HDFS, sqoop, pig, hive, splunk, python and shell to be able to frequently bring in data from various sources and make it available for consumption that can be reused across various projects and usecases.
- Mastery of key development tools such as GIT, build tools, SQL and NOSQL databases, and familiarity with collaboration tools such as Jira and Confluence, data visualization tools such as Tableau, PowerBI, etc. Having strong documentation skills.
- Hands-on technical leadership and support for customer issues and incidents regarding the big data analytic platforms through to resolution. Collaborate with big data and cloud vendors to resolve critical issues, design monitoring & alerting frameworks & performance optimizations.
- Mentoring the team on big data tool and technologies and helping them in all technical aspects including best practices.

### **Data Engineer**

Amazon

06/2015 – 09/2017

Chennai, India

*Projects Handled: Amazon Echo,*

- Hands-on experience with Amazon EC2, Amazon S3, EMR, Athena, Glue, Amazon RDS, VPC, IAM, Amazon Elastic Load Balancing, Auto Scaling, Cloud Front, CloudWatch, SNS, SES, SQS, Kinesis and other services of the AWS family.
- Experience in migration of data across cloud environment to Amazon EC2 clusters . EC2 -to- S3 data synch.
- Creating S3 buckets also managing policies for S3 buckets and Utilized S3 bucket and Glacier for storage and backup on AWS.
- Experience to manage IAM users by creating new users, giving limited access as per needs, assign roles and policies to specific user.
- Familiar with data architecture including data ingestion pipeline design, Hadoop information architecture, data modeling and data mining, machine learning and advanced data processing. Experience optimizing ETL workflows.
- Advanced SQL writing and experience in data mining (SQL, ETL, data warehouse, etc.) and using databases in a business environment with complex datasets
- Developed Python (Pyspark) application to process the peta byte scale log data uploaded by IOT Devices (Amazon Echo).
- Developed Python and Shell scripts to pre-process data and automate manual efforts, which increased productivity of DA by 60%.
- Analysing the millions of log lines from devices to ensure the tuning and logging metrics are up to the mark with help of ELK.

## **EDUCATION**

### **Masters in Business Administration in Database & Systems Management**

University of Madras

01/2016 – 12/2017

Chennai, India

### **Bachelor of Engineering in Computer Science**

REC, Anna University

07/2011 – 05/2015

Vellore, India

## **CERTIFICATES**

Data Engineering with Google Cloud

Participated in Workshop on Machine Learning in IIT Madras.

Certification from Big Data University by IBM on Blockchain Essentials.

Certification from Big Data University by IBM on Big Data, Hadoop, Scala, Spark, Python.