Contact

+919988386679/7973140768(Mobile)

[mohitmnn@gmail.com](mailto:kanikarathore.rr@gmail.com)

www.linkedin.com/in/mohit-mann-aa41a6a5 [(LinkedIn)](https://www.linkedin.com/in/kanika-rathore-59734178?jobid=1234&amp;lipi=urn%3Ali%3Apage%3Ad_jobs_easyapply_pdfgenresume%3Bp4JavGQnSgqS6T59%2Ff53Jg%3D%3D&amp;licu=urn%3Ali%3Acontrol%3Ad_jobs_easyapply_pdfgenresume-v02_profile)

Top Skills Bigdata/Hadoop

ELK

Python/Shell Scripting

AWS Cloud

Languages

English Hindi Punjabi

Certifications

Cloudera Certified Developer for Apache Hadoop(CCDH) certified by Cloudera, 2015.

MapR Certified Cluster Administrator 5.1 (MCCA),2018.

Other Certifications

Developing Hadoop Applications certified by MAPR Academy, 2015

Hadoop Essential certified by MAPR Academy, 2015.

Accessing Hadoop Data Using Hive certified by Big Data University, 2015.

Big Data Fundamentals certified by Big Data University, 2015.

Introduction to Pig certified by Big Data University, 2015.

Building and Monitoring Apache Spark Application certified by MAPR Academy, 2016

Honors-Awards

On The Spot Award for migration of application successfully from DEV to PROD from one network to other network in AWS Cloud.

Participated in Capacity Building Programme on Youth Leadership, Personality Development and Employability organized by RGNIYD at Shimala.

Participated in MozMagnum at Suscet,Tangori.

Participated in an event of Programming Quiz organized by Hackers Club of Suscet at Suscet,Tangori.

Mohit Mann

## Data Engineer | Big data | AWS Cloud | ELK | Spark

## Gurugram, Haryana, India

Summary

* A dynamic and highly motivated graduate student in Computer Science from Punjab Technical University , having good knowledge in the areas Big data Hadoop Technologies , Cloud Platforms, Log analytics tools like ELK Stack
* Proficient in Python, SQL, Shell Scripting
* Have working exposure in Retail, Life-science, telecom, Audit and Retail domain
* Have Good communication skills with effective teamwork and able to quickly adapt & learn new technologies with strong leadership skills.

Work Experience

Nagarro Software Pvt Ltd

Associate Staff Engineer

Sept' 2019 - Present (1.3 year)

Gurgaon, India

**Client Name: US Auditing Company Client**

**Project Name: Audit Work Bench**

**Description: Was part of the Dev team which provides support to Devops team in installing the spark cluster, security of the cluster and optimization of the spark cluster. Help Dev team in application deployment and also involved in writing code as part of application development.**

**Technologies used: Airflow, Python, Apache Spark, ELK, Minio Object storage, Kubernetes, Kafka, ELK**

**Hadoop Distribution: Standalone Spark Cluster**

**Role: Data Engineer and Spark Administration**

**Responsibilities:**

* **Setup Spark Standalone cluster and security of that cluster**
* **Configure HA config for Spark cluster for Automatic Failover**
* **Involved in app deployment from QA to PROD**
* **Involved in the App development activities as per the task provided**
* **Involved in the Issue debugging for the app and then find the possible solution for the same**
* **Capacity planning for the Spark, ELK cluster**
* **Involved in the creating application code in docker image to be deployed on the K8 cluster**

**Client Name:** German Telecom Client

**Project Name:** Anomaly Detection and Fault Tolerance in Telecom Sector

**Description:** Was part of the Data Engineering team whose responsibility was to setup the Platform required for performing the collecting the telecom logs and performing AI/ML. Create data pipelines which will be used to sftp the telecom logs for various network devices. After getting the data from the various network devices, process , prepare the data in the required form which will be used by Data scientist team for ML model development. Also create the end to end pipeline for continuous run of the developed model.

**Technologies used:** ELK, Shell Scripting, Apache Kafka, Apache Spark, Expect Bash scripting, HDFS, SNMP , File-beat

**Hadoop Distribution Used:** Cloudera

**Roles:** Data Engineer

**Responsibilities :**

* Create a Data Engineering Platform which contains Hadoop technologies like HDFS, Spark, Kafka, Hive, Jupyter Notebook, Zeppelin,etc where all the Telecom data is being stored to be used by Data Scientist Team.
* Create Pipelines which will be used to transfer the logs files which includes counter files, syslogs, rpm files from client server to the data engineering platform regularly
* Create the Data engineering and processing pipelines in log-stash component of ELK stack for fetching the ingested logs and processing the log files for providing the input to the Data scientist team
* Responsible for maintaining the Data engineering platform as administrator
* Responsible for Integrating the models developed by Data Scientist team for end to end function as real time
* Responsible for Ingesting the SNMP traps from the network devices to the platform for Anomaly detection.
* Create Scripts for Purging the data time to time.

Tata Consultancy Services Limited

System Engineer

June' 2015 - Aug'2019 (4.3 year)

Noida, India

**Client Name:** US Pharmaceutical Client

**Project Name:** Platform Engineering on cloud Platform

**Duration :** April 2017 - Aug 2019

**Project Description:** Was responsible for building and maintaining the hadoop platform on cloud infrastructure as when required. Also creating automation for application deployment on the hadoop platform. Was part of team with is responsible for exploring new tools for the application team like Grafana, etc. Was also responsible for maintaining the Apache tomcat web-server

**Technologies Used:** AWS cloud Services like EC2, S3, IAM , Hadoop Ecosystem Technologies like HDFS, Hive, Impala, Spark, Sentry , oozie, Kafka, etc, Shell Scripting, Apache tomcat, ELK Stack,

**Hadoop Distribution :** MapR Hadoop Distribution

**Role: Hadoop Administrator and Platform Engineer.**

**Responsibilities:**

* Was responsible for launching on demand hadoop secure or un-secured cluster hosted on cloud based platform for multi-tenant application or single application as per application requirements.
* Was responsible for performance tuning of various ecosystem components like spark, Drill, sentry, impala, etc.
* Was responsible for delivering the secured hadoop platform with client custom security.
* Was responsible for building and maintaining the secure apache-tomcat server and performance tune it.
* Was responsible for deploying the application code on all the environments including production and look into their platform related requirements and automate them using ansible tool.
* Was responsible for the processing of data in hadoop and loading to cloud storage without any platform issues.
* Write Shell scripts that monitors the ingestion process to cloud and give status for that whether that is running or not with is used to populate the dashboard for front end users.
* Responsible for the building the web search and its related activities and issues which is done through elastic search as part of application setup.
* Was responsible for setting the connectivity of MapR cluster with other servers like elastic or web-server.
* Responsible for interoperability testing of hadoop cluster with other tools like talend, alation,redshift,etc

**Client Name:** US Retail Client

**Project Name:** Retail Charts

**Duration :** Sept 2015 - Nov 2016

**Project Description:** Aggregation of data in hadoop so as to get top trending data or reports that are directly delivered to external partners on daily, weekly, monthly and yearly basis.

**Technologies Used:** Hadoop Ecosystem Technologies like HDFS, Hive, oozie, Sqoop, etc, Shell Scripting, PL/SQL

**Hadoop Distribution :** Hortonworks Hadoop Distribution

**Role:** Application Developer

**Responsibilities:**

* Write Hive queries that are used to perform the aggregation of data in hive
* Automation of hive queries through oozie workflows.
* Write Shell scipts that are used to fetch data from hadoop and stores that data in oracle
* I was also involved in making the utility that was used to check the health of hadoop cluster on various parameters and stores the data in oracle tables and then uses that data to verify whether a cluster is working fine or not on timely basis.
* Was responsible for developing code for application and also in optimizing the existing code for better performance.
* Also responsible for writing the automation scripts for application as per requirements

# Skills

|  |  |
| --- | --- |
| Technology Area | Tools |
| Big Data Technology | Hive, Oozie, Map Reduce, Sqoop , Flume, Pig, spark, impala, sentry, Kafka |
| Big Data Cluster Vendors | Hortonworks , Cloudera, MapR, Spark Standalone |
| Operating System | Windows 7,windows 8.1, windows 10,MAC OS X,CentOs |
| Languages | Core Java, Unix shell scripting, , Linux , Basic SQL, Basic Python |
| DataBases | Oracle , MySql |
| Tools/IDE | Eclipse, Android Studio, Net Beans, SQL Developer, HP Service Center - ALM, Pycharm |
| Cloud Platform used | AWS Cloud and its services like EC2, S3, IAM, Cloudwatch |
| Project Management Tools | Jira ,Jira Service Desk, Bit-Bucket, GIT |
| Other Skills | ELK Stack, SNMP Protocol, Minio Storage , Airflow |

### 

### Education

|  |  |  |  |
| --- | --- | --- | --- |
| Period | Degree Title | %/CGP | Institute Name |
| 08/2011-06/2015 | Bachelor of Technology | 78.73% | Institution/ School: Shaheed Udham Singh College Of Engineering & Technology, Tangori |
| 04/2009-03/2011 | XII | 84.4% | Institution/ School: DAV Sr. Sec School, Patiala |
| 04/2001-03/2009 | X | 84% | Institution/ School: Sri Aurobindo International School,Patiala |

Personal Details

Date of Birth: May 31,1993 Sex : Male Nationality : Indian Marital Status: Single Current Location: Gurugram

Passport Details

|  |  |  |  |
| --- | --- | --- | --- |
| **Passport Number** | **Date of Issue** | **Expiry Date** | **Place of Issue** |
| M7188228 | 14-mar-2015 | 13-mar-2025 | Chandigarh, INDI |