|  |  |
| --- | --- |
| **Scala Developer - Noida** | |
| **Candidate Name** | Harshan Rai |
| **Current Company** | IBM |
| **Current Designation** | Senior Technical Services Analyst |
| **Current Location/Native** | Mumbai |
| **Education** | BTech |
| **Experience** | 3.5Years |

|  |  |
| --- | --- |
| **Skills** ( **Scala Developer**) | **Years of Experience** |
| Scala/ Play/Akka/HTTP | 2.5 yrs |
| Scala | 2.5 yrs |
| microservices | 0 yrs |
| MYSQL/ Postgres | 3 yrs |
| Experience in Elastic Search | 0 yrs |
| AWS | 2.6 yrs |
| Experience Git | 2 yrs |
| Ubuntu/ Linux | 3 yrs |

Harshan Rai

+91-8447778856

**[harshan.rai@gmail.com](mailto:harshan.rai@gmail.com)**

# CAREER OBJECTIVE

Seeking a position as a Hadoop & Spark Developer in a good organization which effectively utilizes my expertise in programming and development and where I can prove to be logically strong by delivering exceptional results.

# ABOUT ME

I am a passionate Spark and Hadoop Developer considered a technical expert due to effective skills in developing Shell Scripts, Spark and other Hadoop jobs. Having basic knowledge in OS, Networking and other Computer Science concepts.

# PROFESSIONAL EXPERIENCE

**IBM :** *Senior Technical Services Analyst, DBS Bank Singapore June’20 – Present*

* Recently joined IBM and been allocated to DBS Bank Singapore team.
* Working as a Data Engineer for handling the Data Pipelines and their Spark-Scala codes.
* Trained on new tools as per project requirements such as **Alluxio, Apache Airflow**, **Collibra** etc.
* Handling data for stream processing through Kafka.

**Capgemini :** *Associate Consultant, TechnipFMC July’18 – May’20*

* Meetings with Clients and Managers to discuss and understand all major aspects of a new project, including scope, tasks required and deadlines
* Provided L2/L3 support for all Hadoop related tools with respect to **Data availability**, **Wrangling**, **Refresh frequency updates**, and **Metadata Cataloguing** for our Datalake data to the Data Analysts.
* Worked with Business and Data Analysts to understand their requirements and provide exceptional Automation and Configuration service requests
* Performed data cleaning on unstructured information using **Spark**, **Hive** and wrangling tools like **Trifacta**
* Created and analyzed different **Hive Views and tables** according to the Business requirements
* Used **Bucketed tables** for transferring Postgres tables to Hive with the help of parallel **Sqoop** actions inside OOZIE Workflow
* Contributed Scala code in generating **UCOS-Project Data Pipeline** through **Apache Spark**
* Generated Spark jobs to facilitate data processing by using **Spark-Submit** inside Shell Scripts and scheduled them through **NIFI flows** or **OOZIE Shell/Spark Action**
* Developed highly maintainable **Shell and Scala** codes following the best coding practices
* Automated redundant tasks through **Shell Script and NIFI**
* Scheduled Jobs/Scripts through **Crontab** and **Oozie**
* Metadata Ingestion with their respective Data Classification Tags using **Apache ATLAS Rest APIs**

(GET/POST)

* Ensured that access to Hadoop clusters were secure and private by using Kerberos / Keytab Authentication in NIFI and other tools
* Adopted and Skilled on new technologies to address changing Project and Client needs
* Multi-tasked to keep all assigned tasks and jobs running effectively and efficiently

**Capgemini :** *Senior Analyst, PSA Oct’17 – June’18*

* Designed and modified ETL flows in **Datastage** to provide the required data coming from different raw materials warehouse sites
* Managed the data through **Oracle Client** with the help of **Universe Design Tool (UDT)** and **Information** **Design Tool (IDT)**
* Created **SAP BO Reports** and modified previously built reports
* Handled manufacturing parts inventory data coming from different sources to the Warehouse through Oracle database and ingesting them into different files/BI reports
* Worked on different database tables to daily monitor the data of different sites and reporting it to our Lead and client for inaccuracy or irregularity of data

**INTERNSHIPS AND TRAININGS**

**Capgemini:** *Unix, SQL, ETL, Reporting June’17 – Sept’17*

* Initially, trained on Unix, SQL, ETL Basics and tools like DATASTAGE and SSRS in Capgemini.

**Ducat:** *Android Studio, Training & Project May’16 – July’16*

* Built a Java and Android project in graduation and also made some basic Android Applications.
  + Hadoop Development, Hadoop Admin, Programming, Algorithms, Data Structures

**Key Areas Languages**

**Tools OS**

**TECHNICAL SKILLS**

* + Shell Script, Scala, HIVE, C, Core Java, SQL
  + HDFS, Hive, NIFI, SQOOP, Apache ATLAS, Ranger, Oozie, Spark, Kafka, Alation, Oracle 11g, ETL tools(Pentaho DI, Datastage)
  + REDHAT Linux, Windows
  + Mentorship trough IT Bodhi, Winner in Intra College Singing competition, Under 10 in Mini Marathon

**Scholastic**

**Corporate Hobbies**

**ACHIEVEMENTS AND HOBBIES**

* + Received Rising Star Award’19, Active participation in Cultural Events like Singing
  + Singing, Dancing, Playing Sudoku, 3\*3 Rubik’s Cube, Group Outings

**EDUCATION**

**ABES Engineering College, Ghaziabad** *2013 – 2017*

Bachelor of Technology, Computer Science & Engineering (62.96 %)

**Mahatma Hansraj Modern School, Jhansi** *2011 – 2012*

Central Board of Secondary Education (72.40 %)

**Blue Bells Public School, Jhansi** *2009 – 2010*

Indian Certificate of Secondary Education (73.40 %)