Shravan Kumar

Phone: +1 2026449040

Email:Shravan.hadoopjava@gmail.com

https://www.linkedin.com/in/shravan-kumar-a3a7b9221/

# PROFESSIONAL SYNOPSIS:

* Sincere and competent professional backed up with 14+ years of experience in Machine learning, Data Analytics using big data related technologies, designing, developing, and maintaining applications using Python, Java SE and Java EE technologies.
* Experience in designing, developing, monitoring data pipelines (batch/Streaming) and web and distributed/parallel computing analytic systems using cloudera distribution for Apache Hadoop and Hadoop eco system tools like Spark, MapReduce, Storm, Kafka, Hive, Pig, HDFS and Sqoop.
* Knowledge on Generative AI.
* Working experience on cloud platforms such as Google Cloud, AWS, Databricks and snowflake and ease up development process using tools using as Docker, Kubernetes and CI/CD devops using Jenkins and git.
* Expertise in productionizing machine learning models as APIs and batch models using Kubernetes and cloud.
* Experience in using NoSQL databases like Redis, Cassandra and MongoDB.
* Expertise in application development using JAVA SE, Java EE frameworks such as Spring Framework, Spring Boot, Web Services (SOAP, REST) and ORM technologies such as Hibernate.
* Experience in AWS tools like glue, sage maker, step function, S3, EC2, EMR, ECS, EKS, cloud watch and lambda.
* Experience in GCP toosl like Dataflow, dataproc, app engine, compute engine, Kubernetes engine, big query and pub/sub.

# ACCOMPLISHMENTS:

* **Programming Languages** – Java, Python, Scala.
* **Hadoop Eco System** - MapReduce, Spark, YARN, Sqoop, Hive, Pig, Storm, Kafka, Oozie, Flume, HDFS.
* **Frameworks** – Spring Framework, Spring Boot, Hibernate, Web services (SOAP, REST).
* **UI / Scripting Languages** – HTML, CSS, JavaScript, Node.js, React.js, XML, Shell.
* **Databases** - Oracle 12c, DB2, Postgres.
* **NoSQL databases** - HBase, MongoDB, Redis, Cassandra.
* **Servers** - Apache Tomcat, WebSphere, Jetty.
* **Cloud Platforms** – Google Cloud, Databricks, Snowflake, AWS.
* **Tools** – Airflow, GIT, SVN, Maven, Gradle, Jenkins, Swagger, Apigee, NeoLoad, Gitlab, Docker, Kubernetes, Terraform and Splunk.

# CERTIFICATIONS:

* Databricks certified Data Engineer professional
* AWS Certified Machine Learning Specialty.
* AWS Certified Data Analytics Specialty.
* AWS Certified Solution Architect Professional.
* Google Cloud Certified professional cloud architect.
* Google Cloud Certified Professional data engineer.
* Google Cloud Certified Professional Machine learning Engineer.
* Snowflake Certified SnowPro Core Certified.
* Snowflake Certified SnowPro Advanced Architect.
* Certified Kubernetes Application Developer (CKAD).
* Certified by Oracle as Oracle Certified Professional, Java SE 17 Programmer (OCJP).
* Cloudera Certified Developer for Apache Hadoop (CCDH).
* Certified by CISCO as Network Associate (CCNA).

# EXPERIENCE:

**Client**: NRG Energy  
**Role**: Lead Data Engineer

DATA & AI Platform Jan 2023 – Present  
Built Delta Data Lake for CDP and Data streaming pipelines for GEN AI.

* + **Technologies:** Python, Pyspark, Airflow, Databricks, Delta Lake, Monte Carlo(Data Observability), Azure devops.
  + Build Pyspark jobs to upstream data from its raw format to transform and perform aggregation based on business needs.
  + Orchestrate and entire flow using airflow.
  + Build delta live tables and delta streaming live tables to perform automated change data capture and merging.
  + Designed Delta Lake for databricks to used ACID compliance to files and maintain version history.
  + Designed Data Lake using medallion architecture and implements bronze, silver and gold layers to upstream data from raw to transformed and aggregated ready for analytics.
  + Build Data observability and monitoring using Monte Carlo to perform data validations and detect data drifts.
  + Used azure devops for hosting code as repository, continuous integration and delivery.

**Client**: Merck  
**Role**: Lead java/Big data Developer  
Next Best Action Aug 2020 – Dec 2022  
Models which recommend next best channel for HCP to reach to customer.

* + **Technologies:** Python, java, scala, spark, AWS(Sage maker, Step Function, Lambda, EC2, Cloud Formation, Redshift, Aws Glue, S3), Docker, Terraform, Pyspark, big query.
  + Gather Data from different sources to S3 Bucket.
  + Create required infrastructure using terraform on AWS
  + Integrated AWS QuickSight and tableau to visualize data migration charts and ML model monitoring
  + Cleanse, Transform and store Data in different buckets to render as input to models.
  + Upstream data into different data sources like NoSQL(Cassandra) and SQL(Redshift) databases.
  + Build APIs using Java Spring boot to expose data as service.
  + Build ETL Data pipelines using spark to gather data required for model.
  + Build spark jobs on DataProc to and ingest data to big query.
  + Used pubsub and GCP dataflow(Apache beam) to ingest data to Cassandra.
  + Orchestrate entire sequence using step functions.
  + Improve performance and monitoring of model.

**Client**: State Farm  
**Role**: Lead java/Big data Developer  
Claims Insights Dec 2019 – July 2020

Models which would validates the eligibility of a claim based on Gradient boosting system.

* + **Technologies:** Python, java, Kubernetes, Gitlab, Docker, Postgres, scala, spark, big query.
  + Worked with Data Scientists to understand the data needs for training, validating and productionizing ML models.
  + Create batch and streaming jobs to move required datasets on to Kubernetes PVC.
  + Built ETL pipeline to gather data using Apache spark.
  + Dockerize ML models as images to deploy into Kubernetes as APIs using Python Flask and Batch jobs.
  + Create pipelines to deploy models and run-on test and production Kubernetes environments.
  + Improve performance and monitoring of model.

**Client**: The Home Depot  
**Role**: Lead java/Big data Developer

Merchandizing IT - Product Catalog Dec 2017 – Nov 2019

API and Data Warehouse which holds and provides information about the products sold at Home Depot such as Item Data, SKUs, Vendor/Store info, UPC, GTIN (Global Identification Number), GLN (Global Location Number).

* + **Technologies:** Java 8, Spring Framework, Spring Boot, XML, Google Cloud (Dataflow, Pub-Sub, Big Query, dataproc), Spark, Cassandra, Hibernate, DB2, Pivotal Cloud Foundry, Docker, Kubernetes, Tomcat, NeoLoad, Redis, Postgres.
  + Developed REST APIs for UI to dispense item data, assortment, and vendor information to various consumers.
  + Build data pipelines batch/streaming systems for data migration/transformation from DB2 to Cassandra using Spark/Dataflow (apache beam) and pubsub/Kafka for EDW (Enterprise Data Warehouse) Analytics.
  + Built data validation jobs using dataflow (apache beam) for cleansing and removing discrepancies.
  + Used all major components of google cloud to deploy Apps and Run ETL and Data jobs.
  + Reduced Failures/Errors and latency of APIs using circuit breakers and fall-back policies.
  + Increased fault tolerance of overall systems using chaos engineering principles.

**Client**: American Express  
**Role**: Senior java/Big data Developer

Lending Products July 2014 – Dec 2017

Consists of applications which are built to provide seamless experience to card members in viewing over time limit, transferring balances, enrollments into various promotional offers, request credit line increase and pay partial amount in each billing cycles.

* + **Technologies:** Java 8, Spring Framework, Web services (SOAP, REST), XML, HTML, CSS, JavaScript, React.js, MapReduce, Hive, Pig, HDFS, WebSphere 8.5.
  + Revamped all major applications in portfolio with latest tech stack to provide responsive user experience and improved performance.
  + Exposed All Backend offers and APR data as APIs for various consumers across American Express.
  + Implemented Authentication using Apigee and Spring OAuth.
  + Reduced Response time of APIs by 30% using partial responses and in memory caches.
  + Developed MapReduce jobs to process and analyze huge volume of transactions performed by card members.
  + Designed User Interface for card members using React.js

Credit Suisse AG

Role: java Developer

Syndicate Administration Sept 2013 - June 2014

Primarily used for creating user accounts in Syndicate IOI. This provides an interface for the business users (Internal Syndicate, Sales team) and application administrator to create new users for Syndicate IOI.

* + **Technologies:** HTML, JavaScript, XML, Java SE, Servlets, JSP, SQL, Oracle 10g, WebLogic Server (10.3), JAP (Java Application Platform).
  + Created user experience for different types of users.
  + Designed user clone functionality which saves time in providing privileges to newly created users.
  + Reduced Defects and improved SLA metrics using performance and unit testing.

BASE Education services Pvt.Ltd

**Role**: java Developer

Tuning Fork August 2011 - Aug 2013

Platform connecting students, teachers, parents, franchises, employees, and other players in the co-system of professional education service providers through suite of products. It reduces the turn-around time and efficiency of our day-to-day task, seamlessly scale/expand operations across various centers.

* + **Technologies:** HTML, JSP, JavaScript, XML, Java SE/Java EE, Struts (MVC), Hibernate(ORM), spring (IOC, AOP), Oracle 10g, WebLogic Server (10.3).
  + Created a tool which transforms files in various formats such as excel, word to XML.
  + Implemented MVC and IOC design pattern using Spring framework.
  + Designed user interface and business implementation for various modules.

# Life Care Advisory Unit (LCAU)

**Role**: java Developer

# Electronic Medical Records June 2010 - July 2011

Primarily used for hospital maintenance and divided into 5 modules. Privileges are given based on their role and is useful for capturing patient data and schedule the appointment with Doctor, recording the medicines they prescribed and issues bill

* + **Technologies:** HTML, CSS, XML, JSP, JavaScript, Struts (MVC) Java SE, Servlets, Apache tomcat 6.0, Oracle 10g.
  + Created user interface for various roles.
  + Implemented MVC design architecture for application.
  + Embedded functionality to download reports.
  + Integrated application with payment gateway API for bill payments.

# ACADEMICS:

* + Bachelor of Technology in Electronics and Communication from Jawaharlal Nehru Technological University, Hyderabad, Telangana, India (2006-2010).