

---

### PROFILE SUMMARY

A well-accomplished certified data analyst and Generative AI Engineer with more than 15 years of experience in **Python**, C, C++, Scala, Data Structures, BI, Big-data technologies (Hadoop, Spark, Hive, Spark SQL, Spark-Streaming, Kafka **AI skills (Streamlit, LangChain, Llama2, OpenAI, HuggingFace, RAG, Speech to Text, Vector DB)** and AWS. Experience includes developing Gen-AI applications and maintaining large clusters end to end, data processing and implementation of complex data pipelines, testing of enterprise, mobile application, and middleware component software and seeking a challenging role in your reputed organization where IT skills will be fully utilized.

Developer of Own-Gen-AI-chat (voice based, multi-language support, could train personal and latest events search) <https://voice-based-generative-ai.streamlit.app/>

Papers published on:

- Knowledge to develop application on Streamlets using LLMs and python.
- Multi Language Assistant with user data
- 5 major data structure algorithms that you should know.

### CERTIFICATIONS

- AWS Certified Solutions Architect – Associate
- Confluent Certified Administrator for Apache Kafka
- AWS Certified Bigdata Analytics - Specialty
- Hacker Ranker: Golden badge holder for C++ programming and 5-star award in problem solving skills.
- EICT, IIT Rookie: Big Data Engineering with Hadoop and Spark
- Nano Degree in cloud (AWS) Dev-Ops

### Core Competences

- Strong python programming knowledge (PIL, Pandas, Lambda, matplotlib, NumPy, SciPy, singleton) and Data Structures.
- Worked on developing Generative AI applications
- **Strong knowledge on LangChain, Streamlit and python**
- Hands-on experience of C, Scala, and C++ programming.
- Developed Big Data applications using python, Hadoop, Spark, Kafka architecture and Spark Streaming.
- Hands-on experience of using data structures (List, Queue, Stack, Trees, Tuples, Hash map)
- Worked on deploying and managing application on AWS cloud.
- Hands-on experience of Big Data modules debugging and fixing infra issues.
- Reliable and open-minded with strong team working skills, determined to reach a goal in time as well as the ability to work independently and to prioritize.
- Excellent programming skills and a solid foundation in computer science with strong competencies in Data structures, algorithms, databases, and software design.
- Worked on most of Scala features (case-classes, pattern-matching, Higher-order functions, singleton, companion objects, traits, for-yields, tuples, and multithreading)
- Knowledge and understanding of basic design patterns and strong problem-solving skills & ability to learn in a fast-paced environment.
- Having experience in developing a data pipeline using Kafka to store data into HDFS.
- Worked on review and debugging of java.
- Worked on Docker containers and deployed applications in Kubernetes.
- Experience in ingesting/migrating data into Hadoop (HDFS and Hive) from Apps using Spark and Kafka.
- Worked Cloud Technologies: Oracle cloud and AWS, knowledge of GCP
- Git, GitHub, GitLab, CI/CD pipelines, Docker, Kubernetes, Cloud Formation.
- Business knowledge of Equities, Price-action, Futures and Options. Strong technical and chart analysis skills.
- Performed as a full stack developer including Dev-ops activities.
- significant experience of transitioning On-prem big data platforms to the Cloud

### Corporate-Awards

**Oracle-Star:** for breaking window hack (BI 11g – print issue for a user)

**Oracle-Talent:** for suggesting idea and implementing solution to display unique stack traces for a selected exception.

**Samsung-Knight:** fixing 108 memory leaks and find bug which has impacted complete Tizen OS (C++ null check issue)

## TECHNICAL SKILLS

**Languages:** Python, C, Scala, C++ (Data Structures, STL), SQL

**AI Skills:** Streamlit, LangChain, Llama2, OpenAI, HuggingFace, RAG, Speech to Text, Vector DB

**Cloud Technologies:** Oracle cloud, AWS and GCP

**Dev-Ops:** Git, Git-hub, Git-Lab, AWS Code build, AWS ECR, Docker, Kubernetes.

**Big-Data Technologies:** Hadoop, Spark, RDD, Kafka, Hive, Spark-Streaming

**Mobile Platforms:** SHP, Bada, Linux, WindowsMobile5.0, SLP and Tizen

**Technology:** Multimedia, Content, Mobile Applications, BI, SDK, TDD, BDD, Agile, Scrum, Elastic Search, Kibana

**Tools:** Eclipse, VS2005, VS2010, KlocWork, SVN, git, Clear case, Perforce, PyCharm, IntelliJ, Grafana, Loki, Control Center, Streamsets, Jira

## EDUCATION

**Post Graduate Diploma in** Embedded Systems from CDAC Kochi in 2006

**Bachelor of Technology in** Electronics and Communications from Nagarjuna University in 2005

## WORK HISTORY

**Bank Of Scotland, UK, Generative AI Engineer**

**Dec-2021 to till date.**

**Lloyds Bank, UK, Data Engineer**

**Sep-2021 to Dec-2021**

**Oracle Pvt Ltd, India, Data Analytics Team**

**Jan-2014 to Aug-2021**

**Samsung R&D Operations, India, Multimedia Team**

**Oct-2008 to Jan-2014**

## PROJECT HISTORY

Company: **Bank of Scotland**

Project: Application Hosting

Duration: Apr 2023 to till date

Role: Generative AI Engineer

Environment: Python, Streamlit, LangChain, Vector DB, RAG, Open AI, GPT4All, HuggingFace, Llama2

Docker, GitLab, AWS (Amazon Sage Maker Studio, Bedrock, JumpStart)

GCP (Cloud Run, Vertex AI), PyCharm.

Responsibilities

- Developed multi language supported AI assistant.
- Worked on building Bank-Cora application.
- Worked on developing ask-archie application for with in bank use.
- Created knowledge base of incidents which helped in reducing manual tasks.
- Developed project to upload data and general-purpose search.
- Implemented speech to text and text to speech using google and python libraries.
- Worked on converting text to image using stability diffusion (Hugging face)
- Relevant suggestion: Products and financial advice based on user data.
- Fraud Detection: Proactively identifies suspicious activity for user safety.
- Proactive service: Summarizes user interactions for quicker customer service.
- Personalized marketing: Tailors marketing messages to user interests for better engagement.

Company: **Bank of Scotland**

Project: Streaming Platform

Duration: Dec 2021 to Mar 2023

Role: Sr Data Engineer

Environment: Streamsets, Kafka, HDFS, Spark, Spark-SQL, Kafka Connect, Hive, Git-ops, Grafana, Control Center,

AWS, Loki, Kafka Streaming, Data Collectors, Transformers, Docker, Spark-Streaming, S3,

Jira, Confluence, IntelliJ, Docker, Kubernetes, Bitbucket

Responsibilities

- Implemented program to identify NAP duplicates.
- Design data pipelines on Streamsets and maintain end to end.
- Hermes Regulatory data processing
- Merge/access request approvals, CI/CD pipeline monitoring.
- Monitor BAU of StreamSets prod pipelines.
- Root cause analysis of Inactive jobs/pipelines, data collectors and transformers.
- Implement and debug spark applications (daily transactions, EASRaw jobs)
- updating runtime configurations/resources and jks entries.

- Populating data to s3 bucket using StreamSets pipelines.
- Analyze and define researcher's strategy and determine system architecture and requirement to achieve goals.
- Developed multiple Kafka Producers and Consumers from as per the software requirement specifications.
- Configured Spark streaming to get ongoing information from the Kafka and store the stream information to HDFS.
- Used various spark Transformations and Actions for cleansing the input data.
- Established Kafka cluster and created topics.
- Developed a data pipeline using Kafka to store data into HDFS.
- Knowledge of ACLs and AD groups.
- Knowledge of Kafka connectors, deployment in Kubernetes and debugging the issues.

Company: **Bank of Scotland**

Project: Nano Degree on Cloud Dev-ops

Duration: Sep 2022 to Dec 2022

Role: Sr Data Engineer

Environment: python, IntelliJ, Gitlab, AWS, Docker, Kubernetes, Cloud Formation, AWS code build.

#### Responsibilities

- Deployed Static Website on AWS
- Create S3 bucket and secure via IAM.
- Distribute website via CloudFront.
- Used Groups, Load Balancer, Auto-Scaling to deploy a high-availability web app using CloudFormation.
- Microservices at Scale with AWS and Kubernetes
- Build & run of Docker Images and created an ECR Repository for the Images in AWS Console
- Used AWS Code Build as main CI tool.
- Knowledge of EKS cluster
- Knowledge of using Helm Charts and terraform.
- Used IAM to create roles.
- Used ELB to load web application and Route53 for DNS service.
- Written Lambda function to fetch data from sources.
- Used Cloud-Watch and Cloud-Trail to analyze logs and AWS-api data.
- Fetched data from on premise to AWS using Direct Connect.
- Created a separate VPC and configured subnets (CIDR).

Company: **Lloyds Bank**

Project: IFRS17 Regulation

Duration: Sep 2021 to Dec 2021

Role: Data Engineer

Environment: Scala, Spark, Spark-Streaming, Spark-SQL, HDFS, Hadoop, Hive

#### Responsibilities

- IFRS17 regulatory guidelines implementation.
- Implemented and analyzed Insurance Contracts data.
- Implemented CSM calculation and cashflow generations.
- Data pipelines to prepare data for Measurement models.
- Involved in complete testing process.

Company: **Oracle Pvt Ltd**

Project: Thirdeye-Search

Duration: Feb 2018 to Aug 2021

Role: Data Engineer

Environment: Hadoop, HDFS, Chronos, Spark, Scala, Zookeeper, (OCI) Oracle cloud Infrastructure, Jira, Confluence, Eclipse, Glassfish, Jetty, OHS, Thirdeye-Agents, Elastic Search, Kibana.

#### Responsibilities

- Designed and developed data loading strategies, transformation for business to analyze the datasets.
- Processed flat files in various file formats and stored them as in various partition models in HDFS.
- Responsible for Building, develop, testing shared components that will be used across modules.
- Responsible in performing sort, join, aggregations, filter, and other transformations on the datasets using Spark.
- Responsible in Implementing advanced procedures like text analytics and processing using the in-memory computing capabilities like Spark.
- Experience in extracting appropriate features from data sets to handle bad, null, partial records using Spark SQL.
- Involved in converting Hive/SQL queries into Spark transformations using Spark RDD's.
- Experienced in working with spark eco system using Spark SQL and Scala queries on different formats like Text

file, CSV file.

- Expertized in implementing Spark using Scala and Spark SQL for faster testing and processing of data responsible to manage data from different sources.
- Implemented thirdeye-data pipeline handling mechanism.
- Implemented aha summary and indexed on ELK.
- Implemented Failure and Flaky Analysis algorithms.
- Configured ISL's (Index search server listeners)
- Developed end to end Apache Hadoop cluster from scratch.
- Implemented thirdeye-Collector.
- Worked on configuring Namenode, Datanode, ResourceManager and NodeManagers.
- Worked on migrating Hadoop cluster from on premise to Oracle cloud.
- Extensively involved in debugging Resource manager, Node Manager, Data node and Namenode process.
- Daily monitoring of thirdeye (collector/publisher/ISL) logs to analyze overall stability.
- Analyze Web-Interface and Web-Service logs.
- Analyze cluster matrices to identify faulty nodes.
- Configured OCI instances.
- Configured WebLogic and OHS for security enabling.
- Display indexed data using thirdeye UI.
- Created application load balancers.
- Configured slaves, search-servers, ISL's, Web-Interface hosts.
- Allocated temp usage on boot and worked on block storage plugin/plugin out.

Company: **Oracle Pvt Ltd**

Project: Bug DB Data Processing

Duration: Apr 2017 to Jan 2018

Role: Big Data Developer

Environment: Hadoop, Spark, Hive, Scala, HDFS, Oracle-cloud, Jira, Confluence, Eclipse

#### Responsibilities

- Understand the Big Data Enterprise Architecture.
- Schedule workflows and ingest data into semantic layer based on the business logic and occurrences.
- Migrated and Ingested data into HDFS using and thirdeye agents.
- Developed applications for DTE data handling using Spark, Scala.
- Trouble shoot the Hadoop environment issues like space, Job tracker, prioritize jobs etc.
- Involved in providing support to resolve production issues.
- Coordinated with network team and resolve the source connectivity and BI tools connectivity network issues and firewall port issues.
- Communicate with Hadoop Admin team regarding Space, Resources, permissions etc.

Company: **Oracle Pvt Ltd**

Project: OBIEE Development

Duration: Jan 2014 to Mar 2017

Role: Senior Member of Technical Staff

Environment: C, C++, Data Structures, Visual Studio, Eclipse, OBIEE

#### Responsibilities

- Implementing enhancements, improving the diagnosability and address the customer issues by fixing the code bugs.
- Developed a security feature in downloads module to protect the system from DOS attack.
- Implemented a performance enhancer logic in Ibot to execute the agents faster.
- Responsible for fix/enhance Ibot (Delivers/Agent/Schedulers) issues in the product.
- Responsible for fixing the core dump issues reported by customers in the nqscheduler component in Linux 32/64 and windows 32/64 platforms.

Company: **Samsung R&D**

Project: Tizen Platform Development

Duration: Apr 2009 to Nov 2013

Role: Multimedia Engineer/C++ Developer

Environment: C, C++, Data Structures, Visual Studio, Eclipse, SLP, Bada

#### Responsibilities

- Designed, developed, and tested software applications meeting deadlines and budgets.
- Involved in development of audio and video engines using C++.
- Specified, prototyped, developed, and tested an object-oriented, multiplatform C++ framework containing support to data structures, common algorithms sockets, multi-threading.
- Fixed bugs in the order management system.

- Developed highly efficient algorithms in C++ through both pair-programming and independent work.
- The algorithms improved inventory management benchmarks 18 percent.
- Analyzed code to identify and improve performance bottlenecks and contingencies with shared resources.
- Streamlining the code was instrumental in reducing support costs.
- Significantly contributed to the calculation engine, which was based on evaluation of rules and rule groups.
- Fine-tuned order fulfillment apps that boosted repeat orders and customer loyalty 8 percent.
- Designated “Most Innovative Developer of the Quarter.”
- Written test cases for the developed modules and involved in all phases of testing.
- Ported Audio Out and Audio In modules using ALSA library.
- Media Recorder implementation.
- Developed multi-threaded applications (Audio in and Audio Out)
- Used Bridge design pattern, MVC and Single design pattern.
- C++ STL's are used to implement complex logic in processing media files.
- Implemented Hash Table logic for Tizen platform using C++.

Company: **Samsung R&D**

Project: FldLink Application

Duration: Aug 2006 to Jul 2008

Role: VC++ Developer/ Mobile application developer

Environment: VC++, C, C++, Data Structures, Visual Studio, Eclipse

#### Responsibilities

- Involved in developing complete application from scratch using VC++ and MFC.
- Used different API's of MFC.
- Written test cases using C++ to test the developed modules.
- Actively involved in customer inter-action and fixed the reported bugs.
- Maintained project backup and recovery activities using perforce.
- Analyzed logs to identify potential performance issues and overall application stability.
- Build code using visual studio and resolve all kind of errors.