



Mujahed H.

(Gen AI/Agentic AI Data Cloud Lateral Trainer)

mailto:mujahed.email@gmail.com | +91 00 0000 1234

As an experienced data platform architect and training professional with over **22 years** of experience in online, onsite, and corporate training, I have developed a strong proficiency in analyzing and interpreting large and complex data sets. My ability to discover valuable patterns enables me to provide insightful answers to business questions and identify new opportunities. I am skilled in leveraging machine learning algorithms, cloud computing, visualization, and big data environments to make bold, data-driven decisions. My passion for teaching and training has allowed me to deliver high-quality training sessions that inspire and empower learners to achieve their goals.

SUMMARY

- Passionate about design data architect & crunching System.
- Excellent leadership and organizational skills
- Worked with different Lang Chain language Models (N-gram, statistical, neural, Pretrained, fine-tuning, Evaluation of Language Models)
- Strong architecture, analytical and logical reasoning
- Able to work in deadline-oriented environment
- Designing, implementing, and managing cloud-based big data solutions enables me to impart practical, hands-on knowledge to learners.
- A strong believer in cutting edge technology and providing smart solutions.
- Logically optimized solution through coding and implementation.
- Diligent team player with sound analytical skills of self-defining performance parameters
- Ability to objectively evaluate results.

SKILLS

Programming:

Python:

Language Specifications (Control Flows, Functions, Modules). List comprehensions, lambda, Generators, iterators, context manager, OOP (Multiple Inheritance and Method Resolution Order, Abstract

Base Classes (ABCs) and Interfaces, Operator Overloading and Special Methods, Descriptors and Properties, Class Decorators and Meta classes, Mixin Classes and Composition, Dynamic Class Creation and Modification, Slots and Memory Optimization, Design Patterns and OOP Principles), Mockito,

Package NumPy- Array creations, manipulation, universal functions, operations, iteration, masking filtering, linear algebra operations, masked arrays, memory management, performance optimization like vectorization, indexing, slicing, broadcasting, mathematical functions,

Package Pandas - Series, data frames, indexing, selecting data, data cleaning, manipulation, grouping and aggregating, time series analysis, IO operations- csv, Json, parquet, performance optimization, handling categorical data, reshape & pivotal table, combining data, time zone handling, panda's extensions, adv. Data manipulation, outliers & anomalies parallel processing, debugging & best practices.

Package Matplotlib & Seaborn – creating plots, charts, graphs.

Scikit-learn Package – data handling, supervised learning algorithm, Clustering, PCA, Model Evaluation and Selection, cross validation, metrics for model evaluation, model hyperparameter tuning using grid and random search.

Scala & R:

Classes & Objects, Language Specifications, Closures, Collections, Traits, Patterns, RegEx, Extractors, File IO, Packages, OOPs, Params Types, Exceptions, Map, Annotations, Options, Evaluations, Streaming, Serializations, etc.

Other Programming Languages: JSON, Markdown, YAML, XML.

No Code Platform:

N8N(Node In Node)

Execution & Flow Control: Split In Batches, Merge (Wait/Append Mode), Wait, If, Switch, Error Workflow, Continue On Fail, Execute Workflow, Run Once for Each Item

Scripting & Logic: Function, Function Item, JavaScript, Item Looping, Dynamic Routing, Custom Expressions, Dot Notation, Paired Items, Chained Executions

Data Handling: Binary Data, Base64, Buffer, Data Mapping, Expression Editor, Data Nesting, Flatten JSON, Regex, Templating, JSON.parse/stringify

Authentication & Security: OAuth2, JWT, API Key, Basic Auth, HMAC, Header Validation, Credential Encryption, Environment Variables, Vault

Custom & Extensibility: Custom Nodes, Community Nodes, Node Dev CLI, Custom Modules, Webhook Testing, External APIs, Custom Credentials, TypeScript Node Dev

Deployment & Configuration: Docker, Kubernetes, PostgreSQL, Redis, Worker Mode, Queue Mode, Scaling, Load Balancing, n8n CLI, Webhooks Tunneling, Multi-tenant

Monitoring & Debugging: Execution Logs, Workflow Metrics, Node Timing, Manual Debugging, Custom Error Paths, Trigger Replay, UI Snapshots, Log Streaming

Versioning & Integration: Workflow Versioning, Git Sync, JSON Export, Import via API, n8n API, Persistent Storage, Shared Workflows, Parameter Linking

LangFlow:

Core Concepts: Chains, Agents, Tools, Components, Nodes, Graphs, Flow Canvas, Dynamic Routing

Models & Backends: LLM, OpenAI, Anthropic, Mistral, Ollama, LangChain, Custom Endpoints

Data Handling: Memory, Buffers, Retrieval, Vector Stores, Embeddings, Chroma, FAISS, Weaviate

Custom Logic & Programming: Prompt Templates, Custom Tools, Tool Calling, Function Calling, Python Code Node, Conditionals, Routing Logic

Agent Control: Tool Executor, ReAct, Conversational Agent, Tool Selection, Tool Abstraction, Memory Management, Agent Executor

UI/UX & Visualization: Drag-and-Drop, Node Parameters, Input/Output Preview, Live Debugging, Flow Execution Viewer, Session Tracking

Integration & Extensibility: API Calls, Custom Components, Third-Party Tools, Webhook Integration, LangChain Hub, External APIs

Deployment & Hosting: Docker, Self-Hosting, Streamlit, FastAPI Backend, Local Models, GPU Support

Security & Configs: API Keys, Environment Variables, CORS, Auth Middleware, Custom Config Files

Utilities & Misc: Streaming, Callbacks, Logging, Caching, Batch Processing, Async Support, Flow Saving/Loading

Core & Adv. Machine Learning:

Math & Statistics:

Linear Algebra, derivatives and gradients, distributions and hypothesis testing.

Data Processing & Feature Engineering:

Handling missing data and outliers, Feature selection and extractions, Normalization, Scaling of features.

Supervised Algorithms: Linear Regression, Logistic Regression, Decision Tree, Random Forests, SVM (Support Vector Machine), Naïve Bayes Classifiers, KNN (K-Nearest Neighbors).

Unsupervised Algorithms: Clustering (K-Means, Hierarchical Clustering), Dimensionality reduction Techniques (PCA- Principal Component Analysis, t-SNE)

Reinforcement Learning

Natural Language Processing (NLP):

Topic Modeling: Latent Dirichlet Allocation (LDA), Latent Semantic Analysis (LSA), Non-negative Matrix Factorization (NMF), Hierarchical Dirichlet Process (HDP)

Word Embeddings: Word2Vec, GloVe (Global Vectors for Word Representation), FastText, Doc2Vec

Sequence Labeling: Named Entity Recognition (NER), Part-of-Speech (POS) Tagging, Chunking & Shallow Parsing

Dependency Parsing: Dependency grammar, Transition-based and graph-based parsing algorithms, Parsing techniques like Arc-standard, Arc-eager, and more

Advanced Text Classification: Hierarchical classification, multi-label classification, Ensemble methods for text classification

Sentiment Analysis: Aspect-based sentiment analysis, Fine-grained sentiment analysis, Handling negation and sarcasm

Syntax and Grammar: Context-Free Grammar (CFG), Probabilistic Context-Free Grammar (PCFG), Parsing with CFGs and PCFGs

Machine Translation: Statistical Machine Translation (SMT), Neural Machine Translation (NMT), Phrase-based translation models

Information Extraction: Relation extraction, Event extraction, Template-based extraction

Coreference Resolution: Anaphora and cataphora resolution, Mention detection, Resolution algorithms

Semantics and Ontologies: Semantic Role Labeling (SRL), Frame Semantics, WordNet and other lexical databases, Ontology-based NLP

Dialogue Systems: Intent recognition, Slot filling, Dialog state tracking, Reinforcement learning for dialogue management

Generative Artificial Intelligence (GenAI):

Generative AI, Variational Auto Encoders (VAE), GAN, Auto-regressive model, flow-based model, energy-based model, Text Generation, NLP with GPT-3, MLOps, MLFlow, LLM (H2O Studio, LangChain, GPT4), Probabilistic Modeling, Parametric & Non-parametric Models, Generation and Sampling techniques, Dagshub, Vector Database, Flask Gen AI Model Deployment, DVC.

GenAI Platform (H2O.ai):

In-Memory machine, AutoML, MLOps, GenAI Models (Building, Predicting, Adding Layers), Gradient Boosting Machine Model, Hyperparameter tuning using grid search, XGBoost and Deep Learning Models, handling Time series analysis, NLP with Sentiment analysis experiment setup configuration, etc.

GenAI Platform (Meta LLAMA3, Hugging face):

Architecture text generation methods, optimizing generative models for different applications, Generative Adversarial Network (GAN) in image and video synthesis, Integrating LLMs with other AI Models for multimodal tasks, etc.

Prompt Engineering:

GenAI, LLM, NLP, NER (Named Entity Recognition), Model techniques & Fine tunings, Prompting Strategies, Few & Zero shot learning, Hyperparameter tuning, Future trends & research directions.

GenAI- Generative Artificial Intelligence with Integrations (Application Integrations):

- Microsoft Excel Integrations
 - Data Analysis and Insights
 - Automating Repetitive Tasks
- AI in Presentation Tools (Google Slides/Microsoft PowerPoint)
 - Design Assistance
 - Content Generation
- AI in Email Management
 - Prioritizing and Drafting Responses
- AI in Scheduling and Calendar Management
- AI in Market and Customer Insights
- Analyzing Customer Feedback
- Predictive Analytics
- AI in Content Creation and Marketing
- Strategic Implementation of Generative AI
- Identifying Opportunities for AI Integration

SQL & PL/SQL:

SQL, DML, DDL, SQL Operators, Joins, Sub Queries, Trigger, Cursor, Exceptions, Functions, procedures, Indexes, Sequences, Views & Others Database Objects.

Multi Cloud Technologies:

AWS

Compute: EC2, Elastic Load Balancer, Elastic Beanstalk, Lambda, ECS and EKS.

Storage: EBS, S3, Glacier, EFS, Instance Store, Storage Gateway.

Management: CloudFormation, CloudWatch, Cloud Trail, AWS Auto Scaling.

Networking: VPC, API Gateway, IAM, Security Group, Route53.

Data platform: RDS, NoSQL-DynamoDB, Glue (ETL), AWS Snowflake, Redshift.

Machine Learning: Translate, SageMaker, Comprehend, Lex, Polly, Rekognition, MLOps and MLFlow.

Analytics: Athena, Big Data, EMR (Hadoop, Spark), Kinesis, Quick Sight (BI).

Development: Code Commit, Code Pipeline, Code Whisperer, Cloud9.

Others: AWS Amplify, SNS, SQS and Step Functions.

Azure:

Compute: Virtual Machine.

Storage: Blob Storage.

AI/ML: MLOps, Azure Machine Learning, MLFlow, Azure Databricks for Data Science, Data Lake, Delta Lake, Lake house, Azure ML Pipeline, Azure ML Deployment on AKS.

Big Data: HDInsight, Data Bricks.

Migration: On-Prem to Azure migration, Data Migration Service, Server & Database Migration, SQL Server Azure to Azure Multi Account migration.

Big Data: HDInsight, Data Bricks.

Azure DevOps (VSTS): Azure Boards (Kanban boards), Azure Pipeline (Kubernetes, YAML), Azure Repo (Repo & Github), Azure Test Plans, Azure Artifacts and also deployed Complete End to end Application in Python.

Other Services: Microsoft IIS, Virtual Net (VNet), IAM (Identity and Service Providers), PaaS Services, Azure Active Directory, SSO, OAuth, MFA-Multi Factor Authentication, Directory Service, Domains.

Certification: Microsoft Azure AZ-104, Azure Developer Associate, AI-900, AZ-400 Azure DevOps Engineer, DP-100, DP203.

GCP-Google Cloud Platform:

Compute: Compute Engine Advance Features (Confidential VM, Managed AD, Migrate), Migrate to VM, App Engine GKE and Cloud Functions.

Storage & BigData: Cloud Storage, Cloud SQL, Dataflow, IoT Core, Databricks, Big Query, Spark.

Cloud AI: Speech API, DialogFlow/api.ai (NLU, Chat & Voice Bots).

Apigee Services: Apigee X, Apigee Hybrid, Apigee Edge, Apigee Edge Private Cloud/OPDK, Apigee Micro gateway, Apigee Adapter for Envoy.

Anthos Services: GKE, Anthos Deployment Models (AWS, Azure, Bare Metal, VMware), Anthos Config Management, Cloud Marketplace and Deploy Cloud Run for Anthos, Cloud Volumes by NetApp

Networking: VPC, KMS, Load Balancing, CDN, DNS, Firewall Rules, VPN.

Cloud AIML & APIs: Vertex AI, Vision, Translate, Natural Language, CloudML, Speech API, DialogFlow/api.ai (NLU, Chat & Voice Bots).

Management: Cloud Console, Logging, Debugger & Monitoring.

Security Services: GCP Chronicle SIEM and SOAR,

Containers & Orchestrations: Cloud Build, Cloud Run, Cloud Source Repo, Container Registry, GKE (Istio, GitOps Workflow-ArgoCD, Flux), Cloud Operation Suite.

Other Services: Cloud, IAM, Build, Cloud DataProc- Apache Spark, Data Prep

Containerization: Docker (Dockerfile, AI based Application deployments), AWS ECS, Docker Compose, AKS, ECR, Docker Hub, Docker-Swarm.

Amazon Q Developer: Certified AWS Code Whisper.

Github Co-Pilot: AI-Powered Autocompletion, Code Snippets, Contextual Awareness, Worked Languages (Python, JavaScript, TypeScript), Feedback Iteration, Collaborative Features (pair Programming, Git Integration).

Big Data Ecosystem:

1. Hadoop (Apache-MapReduce, HDFS, Cloudera)
2. Spark (Core, PySpark, SQL Spark, Scala Spark),
3. Databricks (Architecture E2, Runtime, Delta Lake house & Engine, MLFlow)
4. Other BigData Ecosystem (Apache Kafka, HBase, Hive, Cassandra)

ETL & Integrations: Talend, ETL, Data warehouse, MSBI, Business Integration (BI) & Its Testing.

Database: MySQL, MS SQL Server, Oracle, SQLite, RDS, HBase, PostgreSQL, DynamoDB.

Tools: Cloudera Manager, Ambari, StreamSets.

REST API Tools: Postman, JSon, XML, XSLT, XSD.

OS: UNIX, Linux (CentOS, RedHat, Ubuntu), Mac, Windows.

INDUSTRIAL WORKING EXPERIENCE

Client: Facebook, Appatura, NPR (National Population Registration) – Govt. of India, Discovery, MTData (Australia), Department of Arizona (Child Safety), John Hancock, Suez, Goldman Sachs.



Technology Stack:

Machine Learning & AI:

- TensorFlow, PyTorch, H2O.ai, MLFlow, MLOps
- AI/ML Pipelines with Spark, Scikit-learn, XGBoost
- H2O.ai, Spark MLlib, MLFlow (Model Lifecycle Management)

Generative AI & Agentic AI:

- OpenAI GPT-4 / GPT-4o (Foundation Models)
- No Code PaltForms(N8N, LangFlow)
- Meta LLaMA 3, Google Gemini, Anthropic Claude
- HuggingFace Transformers, LangChain / LlamaIndex (for agentic workflows)
- AutoGen / CrewAI (multi-agent systems), RAG (Retrieval-Augmented Generation) Pipelines
- Vector DBs: **Pinecone, Weaviate, FAISS, Chroma**
- Embedding Models: **OpenAI, Cohere, BGE, E5**
- Fine-tuning & Prompt Engineering, Open Source LLMs (e.g., Mistral, Mixtral, Falcon)

Databases & Storage:

- SQL (Advanced level)
- MySQL, PostgreSQL

- Cloud SQL (GCP)
- S3 (AWS), Cloud Storage (GCP)
- Hadoop Ecosystem (HDFS, EMR, etc.)

Cloud Platforms & Services

• Microsoft Azure

- Azure Databricks, Azure Data Factory (ADF), Synapse, Azure DevOps & DataOps
- Azure Data Fabric, Microsoft Cloud Azure (Basic to Advanced), Integration Runtime (Azure & Self-hosted), Triggers, Custom Activities, Error Handling, Monitoring and Logging (Azure Monitor)

• Amazon Web Services (AWS)

- AWS Data Lake Formation, AWS EMR (HBase, Hadoop, Spark)
- AWS Glue, Athena, S3, EC2, AWS Lambda, Containers, QuickSight

• Google Cloud Platform (GCP)

- GCP Data Engineering, Compute Engine, Kubernetes Engine, Cloud Functions
- GCP App Engine, Cloud Storage, Cloud SQL, BigQuery, Pub/Sub, Dataflow, Dataproc
- GCP CCAI, Dialogflow, Translate API, Twilio

Data Engineering & Processing

- Apache Spark Core Engine, PySpark and Spark SQL, Apache Beam, Presto, Drill
- Airflow, Apache NiFi, Kafka, MQTT Protocol, Data Pipelines (ETL/ELT)
- Mapping Data Flows & Visual ETL tools, Data Catalog, Metadata Management
- Data Warehouse Solutions (SQL-based)

CORPORATE TRAINING EXPERIENCE

Client Lists:

- Healthcare Pharmaceuticals Limited (HPL), Bangladesh
- Petroleum Development Oman(PDO), Oman
- Incedo Inc(New Jersey)
- TATA Power
- IBCS Primax

- Code Academy
- Oman Holding International (OHI) Group (IITC)Cognizant
- Capgemini
- John Deere
- Disney
- Zensar (BVW, VPKBIET, KJC, CMR, KCT, Orchid)
- Sandip Foundation
- Nokia
- Jabel Marra
- Incedo
- Great Learning
- BMW Group
- Pepkor Payments & Lending
- Maruti
- IBM(International Business Machine)
- LTI Mindtree
- Shell
- NSE (National Stock Exchange)
- Infosys (EdgeVerve)
- ITC
- Vodafone

Technologies Delivered:

DevOps:

- **CI/CD & Build Tools:** SonarQube, Maven, Jenkins, GitHub Actions, GitLab CI/CD
- **Containerization & Orchestration:** Docker, Docker Compose, Kubernetes (K8s), Helm, Skaffold
- **Cloud-Native Kubernetes Services:** Amazon EKS, Azure AKS, Google GKE
- **Infrastructure as Code (IaC):** Terraform, AWS CloudFormation, Pulumi
- **Monitoring & Logging:** Prometheus, Grafana, ELK Stack (Elasticsearch, Logstash, Kibana), Fluentd

- **Configuration Management:** Ansible, Chef, Puppet

Big Data:

- **Processing Frameworks:** PySpark, Apache Spark, Apache Flink
- **Platforms & Services:** Databricks, Amazon EMR, Google DataProc
- **Orchestration:** Apache Airflow, Azure Data Factory
- **Scalability & Management:** Helm for Kubernetes deployments, Horizontal Pod Autoscaler, Cluster Autoscaler
- **Storage & Lakehouses:** Delta Lake, Apache Hudi, AWS S3, Azure Data Lake Storage (ADLS)
- **Streaming & Messaging:** Kafka, AWS Kinesis, Azure Event Hubs

Security:

- **Cloud Security:** AWS IAM, Azure RBAC, GCP IAM
- **DevSecOps Practices:** Shift-left security, automated security scanning in CI/CD pipelines
- **Security Tools:** HashiCorp Vault, Aqua Security, Snyk, Trivy, AWS Inspector
- **Secrets Management:** AWS Secrets Manager, Azure Key Vault, Kubernetes Secrets
- **Identity & Access Management:** OIDC, SAML, MFA integration

Cloud:

- **AWS Services:** EC2, Lambda, S3, RDS, Glue, EKS, CloudWatch, CloudTrail, IAM, SNS, SQS
- **Azure Services:** Azure Functions, App Services, Azure Blob Storage, Azure SQL, AKS, Azure Monitor
- **GCP Services:** Google Cloud Functions, BigQuery, Cloud Storage, Cloud Pub/Sub, GKE
- **Hybrid/Multi-Cloud:** Anthos (GCP), Azure Arc, AWS Outposts
- **Cloud Automation & Management:** AWS CloudFormation, Azure ARM templates, Google Deployment Manager

AI/ML:

- **Languages & Libraries:** Python, Scikit-learn, TensorFlow, PyTorch, XGBoost, LightGBM
- **Model Development:** Supervised & Unsupervised Learning, NLP (BERT, GPT, spaCy), Time Series Forecasting, Clustering, Feature Engineering

- **Model Serving & Deployment:** MLflow, TensorFlow Serving, Torch-Serve, ONNX, FastAPI
- **MLOps:** Kubeflow, SageMaker Pipelines, TFX, MLflow Tracking, DVC (Data Version Control)

Cloud ML Services:

- **AWS:** SageMaker, Comprehend, Forecast, Rekognition
- **Azure:** Azure ML, Cognitive Services, AutoML
- **GCP:** Vertex AI, AutoML, BigQuery ML
- **Data Labeling & Pipelines:** Amazon SageMaker Ground Truth, Labelbox, Airflow for ML pipelines
- **Experimentation & Monitoring:** Weights & Biases, Neptune.ai, Prometheus for model metrics

Responsible AI: Model explainability (SHAP, LIME), bias detection, fairness audits, differential privacy

RETAIL TRAINING EXPERIENCE

Client: Edureka, Sunbeam (CDAC) (Courses- DAC, DESD, DBDA), APG Learning (Pune)

Technologies:

- AWS Certification Solution Architecture.
- Python Programming for Data Engineers & Architecture
- Apache Spark (PySpark, SQL Spark).
- Shell & PowerShell Scripting, Python Programming
- Azure Data Scientist Associate DP100
- Big Data Platform (Hadoop, Data Bricks, Spark)
- Data Warehousing (Cloud-AWS, Microsoft Azure)
- ETL, Talend, Data Integrations, SQL, PL/SQL.
- Agile Methodology
- Software Testing
- Elasticsearch Kibana, Logstash

MATERIAL EXPERIENCE



Client: Cyforis Inc

Technologies:

- Hands On Lab

- Use Case Preparations
- Recorded Short Sessions on GenAI, Cloud, BigData & DevOps

EDUCATION

- M.C.A (3 Year Post Graduation) – 1stClass, Zakaria Institute of Technology (ZITs)
- B.C.S (Computer Science)– 1stClass, Tom Patrick Information Technology (TPICIT)
- B. Sci. (Bachelor of Science)– Dr. Babasaheb Ambedkar Marathwada University.

CERTIFICATIONS



- Certified GenAI (Generative Artificial Intelligence)
 - Vendor: SAP, Score: 95.70%, Attempt: First Attempt
 - Verification Link: [Click here](#)



- Certified Professional OCI Gen AI
 - Vendor: Oracle University, Status: In-Progress
- Certified AWS Code Whisper (AWS Q Developer)
 - Vendor: Amazon Web Service (AWS), Status: Completed



- Category: Generative AI Developer

- Microsoft Learn
 - Badges: 114+,
 - Level: 11th +
 - Score: 179,125/227,799 XP+
 - Trophies: 23+
 - Transcript: [Click here](#)
- Robotic Process Automation (RPA)
- Certified Adobe Master Collection

PERSONAL DETAILS

Name : Mujahed Hussaini

Passport : S0001234

Languages Known : English, Hindi, Marathi, Arabic and Urdu.