SUMMARY:

- **Data Science Architect** with over **10+ years** of experience in **MLOPS**, **Analytics,** and **Data Science**.
- Proven **track record** of **leading** successful teams and **implementing** innovative solutions to **drive business growth**.
- **Result-oriented** performer, with extensive team **management** and **customer-facing** experience.
- Extensive experience in mapping ad-hoc business **requirements** to a **deliverable analytics** solution.
- **Hands-on experience** on **Cloud** platforms GCP, AWS, Azure, and **DevOps** tools like docker, Kubernetes, and Jenkins.
- Strong **communication** and **leadership** skills with a focus on **collaboration** and **teamwork**.

SKILLS:

Programming Languages	R, Python
Analytics/Data Science	Machine Learning, Neural Networks, Deep Learning, Text Analysis, NLP, Computer Vision, ChatGPT, LLM, GAN
Database	Oracle, SQL, My SQL, Postgres SQL, NoSQL, DynamoDB
Frameworks/Libraries	KERAS, TensorFlow, NLTK, Spacy, MLFlow, Airflow, OpenCV
Cloud	AWS, GCP, Azure, AWS Kinesis, Azure DevOps, Azure Pipelines, Azure Data Factory
DevOps	Docker, Kubernetes, Jenkins, Terraform, IAC, ARM
Visualization Tools	Tableau, Qlik Sense, Plotly Dash
Miscellaneous Tools	GitHub, SVN, Linux, Jira, Visual Studio Code, Jupyter Notebooks, Azure ML Studio, Sagemaker, Databricks

RELEVANT CERTIFICATIONS:

Microsoft Certified: Azure Developer Associate	**Jun 2023**
Google Cloud Associate Engineer	**Dec 2022**
AWS Certified Solutions Architect Professional	**May 2022**
AWS Certified Data Analytics Specialty	**Apr 2022**
Chartered Data Scientist	**Mar 2022**
ATOS Expert Data Intelligence Domain	**Jul 2020**
Google Cloud Platform Professional Data Engineer	**Nov 2019**

PROFESSIONAL EXPERIENCE:

Organization**: EY** Designation**: Senior** **APR 2021-Present**

Project**: Generated AI-powered SQL and Dashboard Generation** Role**: AI Manager** Domain: **Insurance**

- Description:
  - **Leverage** large language models (**LLM**) like **ChatGpt to automate** the generation of SQL **queries** and relevant **dashboards**
- Responsibilities:
  - **Analyzed** source database **schema** and user **requirements** to **design prompts** templates for BI users.
  - Designed **workflows** for users to **interact** with **ChatGpt** to obtain desired accurate **outcomes**.
  - **Architected** the solution using **Azure Cloud** components like Microsoft OpenAI, Azure DevOps, Web Apps, ACR, and ACI.
  - **Led** the **development** and **deployment** of end-to-end solutions on Azure and **demos** to external **clients**.
- Outcome:
  - **Reduced** day-to-day **dependency** of **business** users on the **backend** team by **65%**.
  - **Reduction** in the number of support **tickets** for **data requests** by **60%.**
  - **Lead time** of dashboard **generation reduced** by over **85%**.
- Technical Stack:
  - LLM, ChatGpt, NLP, Postgres DB, Microsoft Azure OpenAI, Azure DevOps, Web Apps, Azure Container Registry, and Azure Container Instance, Plotly Dash

Project**: Retirement Financial Planning Model** Role**: AI Manager** Domain: **Insurance**

- Description:
  - The retirement Financial Planning Model is a **tool** that helps individuals **plan** for their **retirement** by **estimating** their **future** income and expenses and determining how much they need to save to **achieve** their retirement **goals**.
- Responsibilities:
  - **Planned** and **strategized** the **conversion** of the Excel model to the Python **model** using **AI/ML**
  - **Designed** and **deployed** an end-to-end **MLOPS** lifecycle on **Azure**
  - **Managed risk** to ensure that a **high-quality outcome** was **delivered** to the **client**
  - **Provided** technical **expertise** for model development on **Azure ML Studio**
  - **Deployed** a packaged model using **Azure Functions** and **Web Apps**
  - **Handled** the CI/CD cycle using **Azure DevOps**, docker, Azure Kubernetes Service, and **Azure Pipelines**
  - **Deployed infrastructure** components **uniformly** across all **environments** (DEV, UAT, PROD) using **ARM Templates**
- Outcome:
  - **Reduction** in the manual **effort** of developing models by **70%**
  - **Centralized** management of models and access to data increasing **operational efficiency by 60%**
- Technical Stack:
  - Python, Machine Learning, Xgboost, Azure ML Studio, Azure DevOps, Azure Pipeline, Azure Resource Manager, Azure Functions, Web Apps, Azure Container Registry, and Azure Kubernetes Service

Project: Automated DDQ Processing Role**: AI Manager** Domain: **Wealth Management**

- Description:
  - The client was required to provide a **completed** due diligence **questionnaire** to its **customer** on a **quarterly** basis for **compliance** and legal reasons. The goal of the project was to **automate** this **process** for the client.
- Responsibilities:
  - Involved in **Client interaction** on a regular basis to **understand** the **requirements** and **update** on the **progress** report
  - **Implemented** parsing logic in **Python** to extract **components** like text and tables from **docx word** templates.
  - **Trained** ML model to **classify text** from documents into questions, headers, and others.
  - **Developed** a similarity **model** to **match** extracted **questions** with the existing **corpus** to extract **answers** for questions using BERT.
  - **Deployed** end-to-end **model** using **Sagemaker** notebooks on **AWS Step Function**.
  - **Used** schema-less NOSQL **DynamoDB** to save and extract data efficiently.
- Outcome:
  - **Reduction** in **turnaround time** in submitting completed DDQs from **days to minutes**.
  - Yearly **time-effort** savings of **$500k**
- Technical Stack:
  - Python, NLP, Deep Learning, Bert, AWS Sagemaker, AWS Step Functions

Project**: Bias Mitigation in Loan Application – Trusted AI** Role**: Data Science Lead** Domain: **Banking**

- Description:
  - The goal was to **detect and mitigate bias** in **machine learning models** for loan applications as part of the **Trusted AI Initiative**
- Responsibilities:
  - **Researched and experimented** with various statistical and non-statistical **techniques** and **frameworks** to **detect bias** in **outcomes** obtained from machine learning models.
  - Used **AIF-360 library** to perform **constraint optimization** in training **TensorFlow** model to **reduce bias** during training of ml models.
  - Similarly used the **What-If tool** to mitigate bias post-training of ml models.
  - Deployed solution on **Google Kubernetes Engine** using **Docker** and **Jenkins**.
- Outcome:
  - Reduction in bias **leading to fairer outcomes** as well as **reduction of false negatives** by 5**%**.
  - **Reduction of loss of opportunities** for lenders by **7%** as well as **improved interest rates** for nonprivileged groups leading to better financial outcomes by **10%**
- Technical Stack:
  - Python, Machine Learning, Statistics, AIF-360, Google What-If tool, TensorFlow, Docker, Jenkins, Google Kubernetes Engine

Organization: **Atos Syntel** Designation**: Consultant** **May 2017-Apr 2021**

Project**: Intelligent File Ingestion** Role**: NLP Lead** Domain: **Insurance**

- Description:
  - Client received **files** in **ad-hoc formats** from multiple customers.
  - These needed to be **automatically classified and tagged** appropriately to consolidate, organize, and generate reports and statistics.
- Responsibilities:
  - Used spacy **NLP library** in **Python** in **Databricks** notebooks to identify and pre-process files.
  - **Trained and deployed** model using **Azure ML Studio**.
  - **Implemented** the "**Schema/Data Drift**" component to trigger **model retraining and redeployment** if required conditions were met.
  - Built a complete end-to-end **execution pipeline** using **Azure Data Factory**
- Outcome:
  - Automated tagging of files resulting in a reduction of manual effort by 75%.
  - **Real-time** generation of **reports** for the client thus reducing **access time** to obtain **aggregated data** from **days to seconds**.
- Technical Stack:
  - Python, Machine Learning, NLP, Spacy, Databricks Notebooks, Azure ML Studio, Azure Data Factory

Project**: Connected Cars Platform IOT** Role**: Data Science Lead** Domain: **Automotive**

- Description:
  - Driving **telemetric data** from the **unity engine** was captured and fed to Cloud using **AWS IOT Core** which was further used to **predict anomalies and generate alerts** in **real-time** for end users.
- Responsibilities:
  - **Processed** data in **real-time** using **AWS Kinesis Data Streams** and **Amazon Timestream TSDB**.
  - **Trained** model using **Xgboost** and **tracked** model performance using **Mlflow**.
  - Built **data pipelines** using **Airflow** DAG to **detect anomalies** in driving behaviors and **generate alerts** for neighboring vehicles.
  - **Developed and deployed microservices** on containers on **AWS EKS** using **Jenkins** as **CI/CD** tool.
- Outcome:
  - **Consolidation** of the vast amount of **data** on the **cloud** thus **reducing development time** by **70%**.
  - **Real-time** generation of feeds and **alerts** for end users which in turn **reduces the frequency of accidents** by **40%**.
- Technical Stack:
  - Python, Machine Learning, NLP, Spacy, Databricks Notebooks, Azure ML Studio, Azure Data Factory

Project**: Background Removal from I-Card Photo** Role**: Data Scientist** Domain: **Technology**

- Description:
  - Automatic detection and removal of background from photographs submitted for the Identity card.
- Responsibilities:
  - Used **OpenCV** to preprocess image to feed to DL model.
  - Trained **deep learning model** using **TensorFlow** to **detect background** in images**.**
  - Trained **Pix-to-Pix GAN** to **remove background** from photographs.
- Outcome:
  - Improved customer **engagement and satisfaction** by **20%**.
  - Increase in new **customer acquisition** by **10%.**
- Technical Stack:
  - TensorFlow, Generative Adversarial Networks (GAN), Python, OpenCV

Project**: Loan Default Prediction** Role**: Data Scientist** Domain: **Banking**

- Description:
  - The aim of the project was to introduce a **self-learning system** and **reduce false positives**.
- Responsibilities:
  - **Performed** data exploration using Pandas, Numpy, and Tableau.
  - **Preprocessed** data using Python and Pandas
  - Used sklearn for model training, tuning using **hyperparameter** tuning, and **k-fold** cross-validation techniques for model validation.
  - **Created** a **dashboard** for presenting insights using **Tableau**.
- Outcome:
  - Overall, a **20% reduction** was achieved in **false positives** and the system became more **adaptive** to incorporate new patterns.
- Technical Skills:
  - Python, Pandas, Numpy, Tableau, Sklearn

Organization: **3iInfotech** Designation**: Senior Software Developer** **Jan 2015-May 2017**

Project**: SWIFT Message Search Application** Role**: Technical Lead** Domain: **Banking**

- Description:
  - This application was used for users to **view and manage** operations related to **SWIFT messages** used for banking transactions.
- Responsibilities:
  - **Migrated** application from .NET 1.0 to .NET 4.0 MVC framework.
  - **Developed** an application to **generate reports** on **dynamic input** from users.
  - **Enhanced** the **performance** of the application to provide **data in real time**.
  - **Managed** Team to ensure smooth and **efficient delivery** of new features and change requests on time with high-quality delivery.
- Outcome:
  - **Improved performance** of the application to provide data in **real-time** which increased the **efficiency** of **business users by 30%.**
  - **Migrating** to a new framework led to **enhanced security** and **efficient use** of **infrastructure** by **20%**.
- Technical Skills:
  - .NET, C#, MS SQL, MVC

Organization**: Tata Consultancy Services** Designation**: System Engineer** **Dec 2012-Dec 2014**

Project**: Integrated Risk Monitoring System** Role**: Software Developer** Domain: **Finance**

- Description:
  - A major financial entity involved in the transaction of government bonds required an enhanced risk monitoring system.
  - IRMS served as **DataMart** for business users to **view vast amounts of information** and perform risk analytics for the client.
- Responsibilities:
  - **Served** as a core **developer** for the development of a **major ETL process** which was the **backbone** of the application.
  - **Developed** an application to automatically **generate** daily **reports** and **worked** on **improving** the **performance** of **SQL queries**.
- Outcome:
  - Providing customers with real-time data increased client **compliance by 30%** and was one of the **major factors** leading it to obtain a **higher ranking** from authorities thereby **reducing risk** and **increasing** further **business opportunities**.
- Technical Skills:
  - .NET, C#, Oracle

EDUCATION:

MBA – Tech MBA **Jun 2021-May 2023**
Hult International Business School

Executive Post Graduation in Business Analytics and Big Data, Data Science **Jul 2016-Jun 2017**
Aegis School of Business, Data Science & Telecommunication

Bachelor of Engineering Computer Science **Aug 2008-Jun 2012**
Shah and Anchor Kutchi Engineering College, Mumbai University.