# PROFESSIONAL SUMMARY

Senior Data Engineer with over 10 years of experience in Data engineering, Cloud computing, big data processing and Python development and testing across the banking, healthcare, and automotive sectors. A Microsoft Certified Azure Data Engineer, proficient in designing scalable ETL pipelines, data lakes, and cloud-based architectures using Azure, Snowflake, and Databricks. Experienced in Python-based data automation, real-time data processing, and machine learning integration. Adept at SQL optimization, data modeling, and performance tuning, delivering efficient and scalable data solutions.

**SKILLS** & TECHNOLOGIES

* **Programming & Data Processing**: Python, PySpark, SQL, Pandas, NumPy, FastAPI
* **Cloud & Big Data**: Microsoft Azure (Azure Data Factory, Azure Databricks, Azure Synapse), Snowflake, HDFS
* **ETL & Data Pipelines**: SnowSQL, Azure Data factory, Data Warehousing, Databricks
* **Data Modeling & Analytics**: Power BI, Tableau, Data Visualization, Business Intelligence (BI)
* **Automation & CI/CD**: Jenkins, Bitbucket, Git, Robot Framework
* **Machine Learning & AI**: TensorFlow, SciPy, Scikit-learn, Predictive Analytics

# CERTIFICATIONS

* Microsoft Certiﬁed: Azure Data Engineer Associate
* Microsoft Azure Data Fundamentals
* Post Graduate Program in AI and Machine Learning from Great Lakes Executive Learning

# PROFESSIONAL EXPERIENCE

## CAPITAL ONE Henrico, Virginia

**Senior Data Engineer (Contract)JUNE 2024 - Present**

Results-driven Software Engineer with experience in the credit card KYC process. Proficient in developing scalable software solutions, optimizing data pipelines, and ensuring system compliance, ensuring data quality, compliance, and governance. Skilled in using Python, SQL, SnowSQL, and Snowflake to develop Databricks notebooks and build robust reporting systems. Experienced in enhancing fraud detection algorithms and resolving system inefficiencies to support business objectives..

* Analyzed and optimized data pipelines for credit card KYC processes, ensuring compliance and data integrity.
* Developed Databricks notebooks to enhance fraud detection analytics and streamline data processing using PySpark, numpy, pandas, snowflake connectors and matplotlib.
* Generated detailed reports in Databricks notebook that informed strategic business decisions and improved fraud mitigation. Built automated data validation scripts in Python and SnowSQL using databricks, reducing data inconsistencies by 30%.
* Utilized Python and SnowSQL for ETL processes, resolving data quality issues and enhancing reporting accuracy.

## EUROPEAN CENTRAL BANK Germany

**EMPLOYER: T-SYSTEMS PVT. LTD Pune, India**

## Data Engineer (Senior Consultant) March 2022-May 2024

Worked on the company's internal ETL pipeline and data analysis, generated reports using Matplotlib and Seaborn. Responsible for generating monthly, quarterly, semi-annual, and yearly reports using NumPy, Pandas, SciPy, Scikit-learn, and TensorFlow. Developed Python libraries and ETL processes to orchestrate data flow using the Camunda tool and manage data migration to Azure.

* Worked on the company’s internal ETL pipeline, generating reports using Matplotlib, Seaborn, NumPy, Pandas, SciPy, Scikit-learn, and TensorFlow for monthly, quarterly, and yearly analysis.
* Designed and deployed custom Python libraries for data transformation and validation, reducing manual effort in data cleansing, orchestrated data flow with Camunda and managed data migration to Azure, including HDFS systems and Azure Data Factory.
* Led data preprocessing, optimized Python code, and improved pipeline performance in Azure Data Factory and Azure Databricks using PySpark, NumPy, Pandas, pytest, and R.
* Engineered unit tests with pytest and unittest, and migrated data pipelines from Camunda to Azure, ensuring smooth transitions and efficient processing.
* Conducted a proof-of-concept (PoC) on AI models for categorizing sets of values, experimenting with various algorithms to improve predictive accuracy for calculator matrix prediction and internal model building.

## TELUS HEALTHCARE Canada

**EMPLOYER: TELUS INTERNATIONAL Noida, India**

## Senior Software Engineer January 2019 - March2022

Worked on Automaton scripting, contributed towards data access layer development and Data analysis of an application for Hospitals and Doctors to Manage the Patient Treatment information and their Insurance claims all together in one application which replaced the legacy system of Canada’s various Healthcare departments.

* Collaborated with the development team to design and develop a data access layer using Fast API to provide Api endpoints to fetch data for data processing.
* Worked closely with the system development team and manager to disseminate data to various sources after data processing and creating and managing reports after thorough data analysis using power BI.
* Utilised my Python and Selenium skills to develop a robust automation library, empowering the automation team to streamline their processes and enhance overall efficiency using selenium and robot framework in Jenkins.

## AUTOMOTIVE LIGHTING Germany

**EMPLOYER: MAGNETI MARELLI Manesar, India**

## Research and Development Engineer December 2016 - January 2019

The project involved the development of a virtual environment for Automotive Lighting and on road sign detection through camera.. This virtual setup, known as aLTU, utilised Relay and Switch components that could be managed through the Canoe Python library.

* Implemented YOLO Machine learning model to detect images through camera feeds. Deployed ML model and ML performance lineage through MLOps. Used the model output to feed data to ECUs.
* Integrate the necessary functionalities to simulate DTCs within the ALTU simulation framework. This involves utilising the CANoe python llibrary's capabilities to generate and simulate DTCs based on predefined rules and conditions.
* Implement the logic for generating DTCs based on specific conditions, such as sensor readings, system faults, or other predefined triggers that involve writing scripts or algorithms in Python that interact with the CANoe library to generate and simulate DTCs accurately.
* Developed python framework to validate the ALTU simulation and DTC simulation. These packages can be later used to write test cases that should cover various scenarios and edge cases to ensure the accuracy and reliability of the simulation.

## ROYAL BANK OF CANADA Canada

**EMPLOYER: CAPGEMINI Bengaluru, India**

## Senior Software Engineer February 2014 - November 2016

Worked in various projects for Royal Bank of Canada, such as Critical Illness and Anti-Money Laundering. My duties included developing the automation framework from the ground up using Python and the Robot framework.

* Collaborated with the development team, project managers, and stakeholders to understand the software requirements and define the scope of testing. Creating test plans, test cases, and test scripts based on functional and non-functional requirements.
* Designed keywords functions, developing, and maintaining automated test scripts using Python, pytest and utilised Robot Framework, an open-source automation framework, to create test cases, test data, and test libraries.
* Integrated automated tests into Jenkins. Set up and configure Jenkins jobs to execute automated tests regularly and provide feedback on the software's quality.
* Documented and tracked defects using HPQC. Work closely with the development team to provide clear and actionable information about the identified issues and help in their resolution.

## EDUCATION

**RAJASTHAN TECHNICAL UNIVERSITY Rajasthan, India**

***Bachelor of Technology in Electronics and Communication*** **2009-2013**