

Parthasarathy Rajendiran

Senior Python Developer | Machine Learning Engineer

+1-6476732429
parthasarathy.rps@gmail.com
Kitchener, ON, Canada
www.linkedin.com/in/infopartha
<https://github.com/infopartha/>



Summary

Experienced Data Integration Specialist with over **12 years** of expertise in **data warehousing**, **ETL development**, and a strong foundation in **Python** and **advanced analytics**. Proven track record of improving processes through **automation**, enhancing **application performance** and **data quality**, and implementing **scalable data solutions**. Seeking to leverage deep technical knowledge and hands-on experience to contribute to the data management and analytics capabilities.

Education

| | |
|--|------|
| Post Graduate Certificate in Applied Artificial Intelligence and Machine Learning Conestoga College, Ontario, Canada | 2024 |
| Post Graduate Certificate in Predictive Analytics Conestoga College, Ontario, Canada | 2023 |
| Bachelor of Engineering (Computer Science) Anna University, Chennai | 2010 |

Key Skills

| | |
|----------------------------------|---|
| Programming Languages | Python, SQL, Unix Shell scripting, PySpark, R |
| Tools and Packages | Pandas, NumPy, Scikit-learn, TensorFlow, Keras, Scipy, Matplotlib, Seaborn |
| DW Integration Tools & Databases | Informatica PowerCenter, Talend, Teradata, Oracle |
| Bigdata Technologies | HDFS, Hive, Impala, Spark, Azure |
| Concepts and Methods | Data Integration, ETL, Data Warehousing, Data Wrangling, Data Analytics, Machine Learning, Natural Language Processing, Data Science, Predictive Modeling, Feature Engineering, Root Cause Analysis, Performance tuning |

Key Strengths

- **Analytical and Detail-Oriented**
Skilled at conducting thorough data analysis and maintaining acute attention to detail, ensuring high accuracy and reliability in data interpretation and model development.
- **Big-Picture Thinking**
Adept at overseeing projects with a strategic perspective, ensuring that all technical solutions align with larger business objectives and long-term goals.
- **Quick Learner and Adaptable**
Rapidly adapts to new technologies and methodologies, applying them effectively to stay ahead in the dynamic field of AI/ML.

Professional Experience

Senior Python Developer

Ivanti Technology India Pvt. Ltd.

Duration: Sep 2021 - Dec 2022

- Developed data processor and pipeline for patch and integrated patch data from multiple sources using **Python, PySpark, and Docker**.
- Optimized an existing data pipeline by **vectorizing** pandas operations and refactored the logic, reducing the execution time from 7 to 4 hours.
- Implemented a classification model to identify End-of-life products from the scanner plugin reports using **NLP** techniques and **Machine Learning** algorithms.
- Worked on a POC for **deduplicating** product entries using **text distance algorithms**, enhancing **data integrity** across multiple sources.

Cyber Security Works

Duration: Nov 2020 - Aug 2021

- Redesigned data architecture from batch to **real-time processing** and incorporating **multi-threading**, significantly **reducing data latency** from 2 days to 1 day.
- Developed a data validation framework for the new architecture.
- Implemented vulnerability data extraction pipeline for 5 source systems and **containerized data solution** using Docker.
- Prepared a utility to correct/update critical data fields, enabling functional leads and product owners to update and see changes reflected in the system within just 2 hours, significantly enhancing operational efficiency and data accuracy.

Development Team Lead

Cognizant Technology Solutions

Duration: Apr 2020 - Oct 2020

- **Led a team** of 6 developers in developing and deploying credit scoring models built in Python and PySpark, ensuring integration with existing business systems for real-time analytics.
- Supervised **end-to-end delivery** and managed cross-functional (development, testing, DevOps, and functional) teams.
- Developed a logging framework using log4j and a unit testing framework using PyUnit that the whole team can commonly use in their development.
- Established **CI/CD** pipelines using Jenkins and Quality gateway using Sonar.
- Reviewed the team members' code and mentored the team on Python, coding **best practices**, and **agile methodologies**.

Python Developer

Cognizant Technology Solutions

Duration: Mar 2018 - Mar 2020

- Worked on the POCs, which require **Python and PySpark** developments.
- Developed **JSON schema profiler** in Python, which profiles the **nested JSON files** and provides the JSON schema, including the elements, their parents, type of element, and maximum occurrences.
- Spearheaded a **Reverse Engineering** project to identify source-to-target mapping from the raw source files and target files. Worked on data modeling and developed Python scripts to extract and preprocess data from structured and **semi-structured** sources. Employed advanced NLP techniques and ML algorithms for feature extraction, feature engineering, and data wrangling. This project won second place in the Digital Fusion AI Hackathon conducted by Cognizant Digital Business in 2019.
- Developed a **Data Extraction framework** using **PySpark** and **Shell script** to serve the frequently raised data extraction user requests.

- Automated **regular maintenance activities** in the Python server, like the users report and checking the new package installations and their dependencies.
- Developed deletion process in Shell script and hqls, which will perform physical deletion of discontinued data on the major world base tables quarterly.
- Developed **Oozie** workflow, **HQL**, and shell scripts for the monthly extraction of the SOE data and checking its data quality before delivering it downstream.
- Developed a Web application in **Flask, HTML & CSS** for Source Data monitoring POC.
- Developed a dynamic scoring calculation module in Python using pandas and numpy. This generic module dynamically generates calculation formulas based on configuration files.
- Implemented the change data capture process in **shell scripting** using **awk commands**.

ETL Developer

Cognizant Technology Solutions

Duration: May 2011 - Feb 2018

Responsibilities :

- Played offshore **module lead** role and owned the **end-to-end delivery** for many modules.
- Developed many **ETL** and **ELT** jobs in **Informatica PowerCenter** and **Talend**.
- Worked with different kinds of databases and source systems like **Oracle, Teradata, SAP, JDE**, and **flat files** and was involved in data modeling and data profiling.
- Developed Informatica mappings and reusable sessions for **Audit batch control** and Parameter file generation.
- Effectively handled **critical business scenarios** and developed **complex mappings**.
- Written procedures in Oracle and Teradata for loading lookup and dimension tables.
- Extensively worked in Teradata **Pushdown optimization**. Analyzed PDO queries, as well as its advantages and limitations.
- Prepared requirement analysis documents, design documents, source-to-target mapping documents, deployment plan and instructions, requirement traceability matrices, etc.,
- Worked in a **multi-vendor development team** and reported **directly to the client**.
- Worked in **critical production deployments** and **production issues** and resolved issues by coordinating with multiple teams.
- Performed **root cause analysis** and **bug fixing** during SIT and UAT.
- Developed many **Excel VBA-based tools** and **Excel-based dashboards for automating and streamlining day-to-day development activities**, which shortened the turnaround time and improved the quality of work.

Certifications

- Cognizant Certified Data Scientist in 2020.