ANANDKARTHICK

anandkarthickk@gmail.com
www.linkedin.com/in/anandkarthick | California, United States.

Data Geek! Ambitious, Self-Motivated, Optimistic and Intuitive. Qualified professional with over 12 years of experience and outstanding analytical & leadership skills. Always passionate about learning new things and taking on challenges.

Skills & Expertise

- Python, Airflow, Airbyte, DBT, ETL, Spark, Scala, SQL, Collibra, Kafka, Tableau, MPP Systems, docker, Segment, Amplitude, BigEye, Fivetran
- AWS components (including but not limited to S3, EMR, EC2, IAM, Lambda, RDS, Kinesis, Cloudformation)
- Data Integration, Data Analytics, OLAP/OLTP systems, Business Intelligence, Metadata Management, data Modeling, Data Visualization, Data Governance, Data Architecture

Professional Summary

Senior Software Engineering - Data
Education | **Udacity**
SF Bay Area, CA | Sep 2021 - Present

- Designed and Implemented **framework** to **govern** events data from customer data platform using **Python**, **DBT**, **Airflow,** and third party **API**.
- Designed and created self serving **tracking plan** in **Tableau** & **governance** dashboard with key metrics, access insights and issues enabling **Product** and **Engineering teams** to **proactively identify issues** and avoid redundancies.
- **Designed** Central **Metadata** repository with **Apache Atlas** and built **data pipelines** to ingest/extract data for operational reporting & dashboards enabling the org with cleaner data catalog.
- **Contribution** to **open source** : **Airbyte** connector for Clari to bring in sales/revenue data, **bug fix** in **Apache Atlas Client** (Metadata management).
- Built robust **data pipelines in Airflow** with automated **unit test** cases, **data modeling** to integrate data via API from third party sources to **data lake**, operators for Airflow, dbt models to enable analytics on student retention data.
- Built **reusable data masking** and **encryption** frameworks for **PII data**, reusable tools for **audit compliance** and **handling**.

Data Engineering
Financial Services | **Accenture**
Entry: Analyst & Exit : Associate Manager
India | Connecticut
Jun 2010 - Sep 2021

- **Lead data engineering efforts** and building scalable data pipelines, cloud apps to integrate data from internal and third party systems.
- Built **reusable data frameworks** and **integrations** tools using **Spark/Hadoop** that integrates external data to create different data layers to cater **behavioral analytics** and data science teams.
- Built **Spark testing framework** to enable test data management and **reduce** deployment lead time **by 30%.**
- Built basic **classification models**, **KPI driven metrics framework** using **correlation** analysis worked with **NLTK**, **spaCy**, data discovery for **data science** projects to cater different functional and technical teams.
- **Automated** existing business process **reducing 40%** of processing time, **25% increase** in business **efficiency** to bring new business.
- **Built** a rule based **framework** to validate business data **reducing** data validation time **by 50%**. Optimized various ETL pipelines to **improve** response time and batch processing time **by 25%**.
- **Reduced** data issues and production bugs (~**60%**) with efficient data quality checks, automated data validation.
- ETL **development** to build **data marts** for various client campaigns **enabling** daily/weekly **reporting** to **improve** customer targeting and **operational efficiency**. Implemented automation to improve batch performance **by 30%** and **SLA by 6%**.
- **Developed** and **Implemented** test automation framework for data integration projects to **improve** deployment time **by 25%**. Built code review, testing utilities to improve **deployment lead time by 40%**.

Certifications

- INS 21 – Introduction to Property & Casualty Insurance @ AICPU
- The Data Scientist's Toolbox @ Coursera
- Scaled Agile Framework® Agilist (SA) Certification Program @ Scaled Agile
- Technology Architecture @ Accenture
- Python for Data Science @ UCSanDiegoX
- Probability & Statistics in Data Science using Python @ UCSanDiegoX
- Machine Learning Fundamentals @ UCSanDiegoX
- Python Basics, SQL Basics (Hackerrank)

Education

B.Tech in Information & Communication Technology from SASTRA University (India)

External Projects

Link to GitHub : https://github.com/Anandkarthick

Public events/Press

NERD Summit 2018 (Speaker) : Data Science with Python

https://2018.nerdsummit.org/ (Saturday 2 pm)

Meetup - Greater Hartford Python (Speaker)

https://www.meetup.com/greaterhartfordpython/events/249239821/