Abhishek Kumar
Email: abhishek7277905904@gmail.com
Mobile: 7277905904

Professional summary:

- Having a total of **7.9 plus years** of experience as ETL developer and multiple technologies..

Professional Experience:

- Working as **Data Engineer** for **Biofourmis, Bangalore** from **October 2020** till **July 6th,2023.**.
- Worked as **Software Engineer** for **maplelabs.com, Bangalore** from **September 2019** till **September 2020**.
- Worked as **Application Development Analyst** for **Accenture, Bangalore** from **Nov 2015** till **September 2019**.

Technical skills:

Primary Skills	**Data Engineering, Python, AWS Redshift**
Secondary Skills	Tableau, REST API, Django, Flask, NLP, Applied ML ,PostgreSQL, Hbase, Elastic search, Dynamodb, Redis
Web Servers	Nginx
Operating Systems	Linux
Development Tools Used	Docker, GIT, VScode

Company : Biofourmis Inc.

Project Profile #7: [October,2020 - July 6th,2023]

Project /Role	**Data Warehouse Engineer (ETL engineer, Redshift SQL developer)**
Technologies	**AWS Redshift, S3, Airflow, Python, Pyspark, Fission function, Kubernetes, Mongodb, Pytest, AWS lambda, Celery, Flask, Kafka, Tableau**

Project Description: The project involved creating a real time data streaming platform by using Pyspark and also using the serverless function and deploying them over the EKS cluster.

Responsibilities:

- Developed ETL pipelines for several data sources comprising databases, SQL and NoSQL and file based.
- Understanding the requirements for the current business problem and applying it over Pyspark for a near real time data streaming platform for medical devices and sensor devices.
- Developing the serverless function with k8s fission function as well as lambda.
- Created the data pipeline for the new sensors onboarding the SAAS platform, integrating them and testing them end-to-end.
- Worked over data science rule-based algorithms regarding scheduling which alerts the medical professional over increased threshold.
- Implemented the design pattern for code scalability and followed and mentored for best code practices.
- Resolving production issues based on memory and large data and airflow issues.

Company : MapleLabs Solutions Private Limited

Project Profile #6:[May 2019 - Sept 2019][5 months]

Name of the project	Snappyflow Poller
Technologies	Python, Django, Next.js (React.js framework), AWS, Nginx, Redux
Responsibilities	Understanding the requirements for the current business problem as well as the upstream application and building the application to visualize data and configuration related to project, endpoint and plugin information present in AWS and AZURE based on tags.

Project Profile #5: [Sept,2019 - May,2019][9 month]

Name of the project	Data Capital Management Web Service (Software Engineering: Data)
Client	Data Capital Management (US based)
Technologies	Python, S3, SQS, Redis, Redis cache, EFS, Flask, Pandas, RDS, AWS

Project Description: The project involved converting the standalone architecture to a microservice architecture using python and AWS services. The client's domain was Investment strategy.

Responsibilities:

● Understanding the requirements for the current business problem and applying the architectural decisions made with the client in the design phase.

● Developed the architecture end-to-end using the system design principles along with Redis cache at the client side as well as on the server side.

● Worked over the decision making of the bottlenecks faced during the development phase using multiprocessing, pickling of data, and exposing the services using REST APIs.

Company : Accenture Services Pvt Ltd

Project Profile #4:[March 2018- Sept 2019] [7 months]

Name of the project	: Data Science Labs (MLOps)
Client	: Ericsson
Technologies	: Django, Python, Jupyter notebook, Docker, R-language, RPlumber, Pandas library, scikit learn
Responsibilities	● Understanding the requirements for the current business problems and applying it via Python and R and deploying it as an API with the help of Django and RPlumber in Docker.

	• Applied Random forest classifier on one use case which involved classifying customers as SOR or Non-SOR. • Worked on time-series analysis using R library for finding the anomalies and deploying with RPlumber. • MLOps

Project Profile #3:[July, 2017 - Feb 2018][18 months]

Name of the project	IDAG (GDPR as a Service), **Data Engineering**
Client	: Accenture Internal Project
Technologies **Libraries used** **NoSQL database** **Resposibilities**	: Django, Python, Jupyter notebook, Docker, Pyspark, Swagger : Scikit-learn, pandas : HBase Responsible for building pipelines for data governance flows, flagging PII values present in the database with help of pyspark and python

Project Profile #2: [Feb 2016 - July 2017][6 months]

Name of the project	DIS (Data intelligence Suite)
Client	Accenture Internal Project
Technologies **Resposibilities:**	Python, Flask, REST API, HBase, scikit-learn, pandas, Docker Responsible for creating the APIs for the drag and drop and other features required for pulling the data from the client's database and applying ML algos on top of it.

.

Project Profile #1: [Jan 2015, Jan 2016][12 months]

Name of the project : GDSS
Client : NBN Co (Australia)
Technologies : Flask, Python, PostgreSQL, REST API, AWS **Responsibilities**: Responsible for creating python and flask based APIs for the data processing of huge geo location files and loading them in postgres database.

Qualification: B.E. from Dayananda Sagar College of Engineering, Bangalore, graduated in 2015.

Certification: Completed Deeplearning.ai specialization from Coursera.org in 2019.