

## SUMMARY

- > 5+ years of IT industry experience. (5+ years of work experience + 6 months of internship)
- > Technical proficiency and intensive experience in OO Python. This includes generic modules to application, backend to web pages, communicating with external interfaces from the application, APIs from application to external entity, Parsing JSON and logs.
- > Expertise in scraping of data from files(PDF, excel, csv, word, etc.), website and APIs.
- > Expertise in design and development of various web and enterprise applications using technologies like OO Python, JSON, and Web Services.
- > UI layer development using technologies like HTML, CSS and JS.
- > Thorough knowledge of UNIX/LINUX.
- > Database Knowledge using SQL and Couch DB.
- > Experience using AWS cloud services - AWS EC2, AWS S3 buckets, AWS Cloud Formation, AWS VPC.
- > Experience using Azure cloud services – Azure Blob Storage, Azure Data Factory, Azure Synapse, Azure Databricks, Azure Logic App, Azure DevOps.
- > Experience using Google cloud services - Google Function, Google Cloud Storage, Google Cloud Scheduler, Google Cloud BigQuery.
- > Designed Functional Specification Document, Technical Design Document, Release Document and Code Documentation for future references.
- > Good communication skills, analytical, problem solving, leading and learning skills as well as keen interest in the emerging technologies, excellent team player.
- ❖ Having B1/B2 visa to USA valid till 2027.

## PERSONAL SKILLS

- Technologies : Big Data, AWS, Azure, Python, Numpy, Pandas, Dask, Ruby, Flask, Boto, Boto3, JS, HTML, CSS, JIRA, VSTS, Airflow, Docker, Azure DevOps.
- Operating Systems : Microsoft Windows (XP, 7, 8), Linux.
- Personal : Strong work ethic, Teamwork, Perseverance.

## PROFESSIONAL EXPERIENCE – Internship & projects

**06-2018 to Present**

**Shell India Markets Pvt Ltd.**

**Data Engineer**

**Title:** Evergreening the Pipelines

**Technologies:** Python, SQL Stored Procedures, Databricks, API, Azure DevOps, Azure – Data Factory, Synapse.

**Description:** Evergreening the pipeline is to validate the Data Quality Check, API validate check, Timeliness Check and PDF Structure Check. Further notify in case of any failure in to the table which will be monitored in Power bi.

**Title:** WorldScale - Scraping Circulars Pipeline

**Technologies:** Python, Tabula, Databricks, Azure DevOps, SQL Stored Procedures, Azure – Blob Storage, Data Factory, Synapse.

**Description:** World Scale - Scraping Circulars Pipeline is to extract data from PDF and store them into Azure Blob Storage. Further this data is stored into Azure Datawarehouse (Synapse) in Multiple layers – RAW, Staging, Curated and Reporting. Further these are fed into Power Bi.

**Title:** WorldScale - Scraping Flat rates from PDF Pipeline

**Technologies:** Python, Tabula, Databricks, Azure DevOps, SQL Stored Procedures, Azure – Blob Storage, Data Factory, Synapse.

**Description:** World Scale - Scraping Flat rates from PDF Pipeline is to extract data from PDF and store them into Azure Blob Storage. Further this data is stored into Azure Datawarehouse (Synapse) in Multiple layers – RAW, Staging, Curated and Reporting. Further these are fed into Power Bi.

**Title:** WorldScale - Scraping Flat rates from website Pipeline

**Technologies:** Python, Databricks, Azure DevOps, SQL Stored Procedures, Azure – Blob Storage, Data Factory, Synapse.

**Description:** World Scale - Scraping Flat rates from website Pipeline is to extract data from worldscale website, where we have to automatically login to the website via script and send queries and scrape the required data and store them into Azure Blob Storage. Further this data is stored into Azure Datawarehouse (Synapse) in Multiple layers – RAW, Staging, Curated and Reporting. Further these are fed into Power Bi.

**Title:** Moe - Google Mobility Data Pipeline

**Technologies:** Python, Databricks, Azure DevOps, SQL Stored Procedures, Azure – Blob Storage, Data Factory, Synapse.

**Description:** Google Mobility Data Pipeline is to extract csv data on daily basis from google mobility website and store them into Azure Blob Storage. Further this data is stored into Azure Datawarehouse (Synapse) in Multiple layers – RAW, Staging and Reporting. Further these are fed into Power Bi.

**Title:** Moe - Apple Data Pipeline

**Technologies:** Python, Databricks, Azure DevOps, SQL Stored Procedures, Azure – Blob Storage, Data Factory, Synapse.

**Description:** Apple Data Pipeline is to extract csv data on daily basis from apple mobility website and store them into Azure Blob Storage. Further this data is stored into Azure Datawarehouse (Synapse) in Multiple layers – RAW, Staging and Reporting. Further these are fed into Power Bi.

**Title:** Moe - Retail Sell IN Data Pipeline

**Technologies:** Python, Databricks, GSAP, SAP WEBI, Azure DevOps, SQL Stored Procedures, Azure – Blob Storage, Data Factory, Synapse.

**Description:** Retail Sell IN Data Pipeline is to Build an automated pipeline to Extract Data from GSAP to Azure Storage. Further process the data into required format to store them into DB Tables. Further this data is stored into Azure Datawarehouse (Synapse) in Multiple layers – RAW, Delta, Staging and Reporting. Further these are fed into Power Bi.

**Title:** Moe - Calendar Data

**Technologies:** Python, Databricks, Azure DevOps, Azure – Blob Storage, Synapse.

**Description:** Calendar Data is to generate 2019 and 2020 Calendar with following details -- Year, Quarter, Month, Days in month, Date, Day, Day of week, Day of year, Week Number. Also generate list of holidays for 28 countries for 2019 and 2020 from a website. And scrape the countries which has date of initial lockdown in the given file. Further this data is stored into Azure Datawarehouse (Synapse) as Reference tables. Further these are fed into Power Bi.

**Title:** Moe - Retail Sell IN China and Hong-Kong Data Pipeline

**Technologies:** Python, Email Services, Databricks, Azure DevOps, SQL Stored Procedures, Azure – Blob Storage, Data Factory, Synapse, LogicApp.

**Description:** Retail Sell IN China and Hong-Kong Data Pipeline is to extract xlsx data on daily basis from Email and store them into Azure Blob Storage in the form of csv. Further this data is stored into Azure Datawarehouse (Synapse) in Multiple layers – RAW, Staging and Reporting. Further these are fed into Power Bi.

**Title:** Moe - Tomtom Traffic Congestion Pipeline

**Technologies:** Python, Email Services, Databricks, Azure DevOps, SQL Stored Procedures, Azure – Blob Storage, Data Factory, Synapse.

**Description:** Tomtom Traffic Congestion Pipeline is to extract Daily Traffic based on city, City Ranking based on congestion, Weekly and Yearly Difference compared to current day for each city in form JSON data on daily basis from APIs of Tomtom website and store them into Azure Blob Storage in the form of csv. And also Generate an excel report with Analysis and share it via email to the Trade Desk. Further this data is stored into Azure Datawarehouse (Synapse) in Multiple layers – RAW, Staging and Reporting. And we also share the data into Shared Drive for other teams to use the data. Further these are fed into Power BI.

**Title:** Moe - Nigerian Daily Stocks Pipeline

**Technologies:** Python, Tabula, Databricks, Airflow, Azure DevOps, SQL Stored Procedures, Azure – Blob Storage, Data Factory, Synapse.

**Description:** Nigerian Daily Stocks out Pipeline is to extract data from PDF on daily basis from APIs of Nigeria PPPRA government website and store them into Azure Blob Storage in the form of csv. Also, send an email with extracted csv attachment to users. Further this data is stored into Azure Datawarehouse (Synapse) in Multiple layers – RAW, Staging and Reporting. Further these are fed into Power BI.

**Title:** Moe - Nigerian Daily Truck out Pipeline

**Technologies:** Python, Tabula, Databricks, Airflow, Azure DevOps, SQL Stored Procedures, Azure – Blob Storage, Data Factory, Synapse.

**Description:** Nigerian Daily Truck out Pipeline is to extract data from PDF on daily basis from APIs of Nigeria PPPRA government website and store them into Azure Blob Storage in the form of csv. Also, send an email with extracted csv attachment to users. Further this data is stored into Azure Datawarehouse (Synapse) in Multiple layers – RAW, Staging and Reporting. Further these are fed into Power BI.

**Title:** Loading US Shell Factual and Vantage Data into Datawarehouse

**Technologies:** Python, Dask, SQL Stored Procedures, Azure – Blob Storage, Synapse.

**Description:** Extract Data from Shared Drive and convert them into csv format and store them into multiple files as it should not exceed 500k data lines. Further it is stored into Azure Blob Storage and Azure Datawarehouse (Synapse).

**Title:** Loading data related to Aviation as Covid19 impact in January 2020 into Datawarehouse

**Technologies:** Python, Dask, Azure – Blob Storage, Synapse.

**Description:** Extract Data from Shared Drive and convert them into csv format and store them into multiple files as it should not exceed 500k data lines. Further it is stored into Azure Blob Storage and Azure Datawarehouse (Synapse).

**Title:** Mario - Singapore API Pipeline

**Technologies:** Python, Pandas, API, JSON, Docker, GitHub, GC – Storage, Functions, Scheduler, BigQuery.

**Description:** Singapore API Pipeline is to extract JSON data on daily basis from multiple APIs of Singapore MPA website and store them into Google Cloud Storage in the form of csv. Further this data is stored into Google Cloud BigQuery. Further these are used by Data Scientists.

---

**Software Engineer**

**Title:** ARIT

**Technologies:** SAP UI5 (HTML, Xml, JS).

**Description:** ARIT Time writing EPF is a UI application to monitor the Time Recorded by the Employee and to generate the reports for Business analysis.

**Title:** US Audit

**Technologies:** SAP UI5 (HTML, Xml, JS), Python (NymPy, Pandas, Matplotlib).

**Description:** Automation of Tax filing process in US region.

**Title:** GTIL

**Technologies:** SAP UI5 (HTML, Xml, JS).

**Description:** GTIL (Global Tax Integration Layer) is a Tool with multiple application which is used for automating Tax filing Process around the Globe.

**Title:** GTIL

**Technologies:** SAP UI5 (HTML, Xml, JS).

**Description:** GTIL (Global Tax Integration Layer) is a Tool with multiple application which is used for automating Tax filing Process around the Globe.

---

## **11-2015 to 05-2018**

## **ASM Technologies Ltd.**

**Software Engineer**

**Client:** CloudPassage

**Title:** Halo (R38, S13), (R39, S14), (R40, S15), (R41, S16) & (R42, S17).

**Technologies:** Python, Ruby, AWS.

**Description:** CloudPassage Halo provides instant visibility and continuous protection for servers in any combination of data centers, private clouds and public clouds. The Halo platform is delivered as a service, so it deploys in minutes and scales on-demand. Halo uses minimal system resources, so layered security can be deployed where it counts, right at every workload – servers, instances and containers.

Halo is the only platform purpose-built for broadly automated, portable, scalable, on-demand security and compliance.

**Title:** Test Management System

**Technologies:** Ruby, Jenkins, Bootstrap. JS, HTML, PSQL.

**Description:** Representing the Automated Test Results on different grids and comparing the latest and previous results. Representing all manual use cases based on features along with their result. And features to add, search and edit manual use cases.

---

**Client:** FourV Systems

**Title:** GreySpark Cyber Security.

**Technologies:** Big Data, AWS, Python, Flask, Boto, Boto3, SQL, Angular JS, HTML, Highcharts, Jenkins.

**Description:** Modern businesses use data to make more informed decisions. FourV Systems use machine learning and statistical analytics to calculate Cyber Risk Indicators from existing system and security sensor data in organizations to identify quantify and manage IT risk. They lift the fog from cyber-security operations performance, calculating performance indicators of actual risk, mapping these to controls and value, and giving technical and business leadership the ability to accept, manage or transfer cyber risk – the same as any other risk.

GreySpark identifies and measures risk factors by normalizing and evaluating threat, vulnerability, and operational events from an organization's sensors and systems. The calculated risk factors drive simple risk indices and provide a continuous view of IT security risk trends and IT security operation's performance.

GreySpark Cyber integrates with workflow management systems in order to automate the assignment of investigation and resolution tasks.

---

**Client:** ONF

**Title:** Flow Manager UI.

**Technologies:** Raspberry Pi, Zodiac FX, Python, Flask, Couch DB, Angular JS, HTML and D3 (Force-Directed Graph).

**Description:** It reads the real-time information of flows on the switch managed by a controller and it is difficult to efficiently manage the flows in a switch. So we have developed the User Interface to represent the switch's data in legible and readable tables. The UI interface is enhanced with several functionalities like View, Search, and Delete in the interface.

---

**Client:** Ruckus

**Title:** Wireless Monitoring Server (V2) and (V3).

**Technologies:** Python, Flask, SQL, Angular JS, HTML, Highcharts and D3 (Force-Directed Graph).

**Description:** Ruckus Wireless Monitoring System which is used in monitoring the status and activities of the Zone Directors, SCGs, APs and Clients.

This system enables the user to add and monitor Zone Directors/SCGs. The user can track the number of Access Points (APs) and the number of client devices, Zone Data Transmission associated with the Zone Directors/SCGs, Busiest APs per device. System can able to monitor particular APs like failover, reboot and Rx/Tx transmitted, Client count and send status of that APs.

It can notify the users about the status of the Zone Directors/SCGs through email and SMS. It will send alerts on failover, status change and version change. It will also send reports (graphs), Historical Logs through email to the users. The device details, client details, and logs which are collected using CLI/ SNMP/SYSLOG/JSON automatically per polling interval.

---

**6-2014 to 11-2014**

**Ciber-Sites India Pvt Ltd.**

---

**Trainee-SAP ABAP Consultant**

**Title:** Job Controller- Automation of Background Jobs.

**Tools:** ERP 6.0 [SD and MM] (SE38, SE11, SM30, SE37, SM36 & SM37)

**Language:** ABAP

**Description:** Job Controller in an application where in which we create Background Jobs and schedule them through ABAP Program extracting the required data from given ZTable. These Background Jobs perform the particular user requirements like generating reports etc., in the background even in users' absence.

---

**02-2013 to 04-2013**

**HAL-Aerospace – Bangalore.**

---

**Trainee - HAL-Aerospace Division**

**Title:** Desktop Sharing/Remote Desktop Monitoring through Secure Network.

**Tools:** Eclipse

**Language:** JAVA

**Description:** Application Sharing and Desktop Sharing allows sharing any application with one or more people over the LAN. The participants receive the screen-view of the shared application from the server. Their mouse and keyboard events are delivered and regenerated at the server. Application Sharing and Desktop Sharing enables collaborative work, software tutoring and e-learning over the LAN. We have developed an application and desktop sharing platform called ADS which is efficient, reliable, operating system independent, scales well, supports all applications and features true application sharing. It also helps the administrator to monitor his subordinates work by monitoring their desktops. As a proof of monitoring, it captures periodic screen shots and provides video recording whenever the administrator wishes to record the work done by his subordinates.

---

**Rajarajeswari College of Engineering, Bangalore.**

---

**Title:** Accident Avoidance in Vehicles by Implementation of Driver Drowsiness Detection, Black Box and Obstacle Detection.

**Tools:** MAT Lab, Netbeans and Arduino

**Language:** MAT Lab, Arduino Sketch.

**Description:** Drivers when driving long separations without general breaks run a high danger of getting to be drained. A state where even the specialists neglect to remember it sufficiently early. Review tells that around 25 percent of all genuine motorway mishaps happen because of languid drivers who need a rest. Driver weariness causes more street mischances than beverage driving. Consideration help can caution of distractedness and laziness of the driver in an amplified velocity range. What's more, it informs driver's ebb and flow condition of weariness and the driving time subsequent to the last break. A minimal effort framework which gives answer for the current car control issues. This framework has two primary rule segments. One is Vehicle to Vehicle Collision Avoidance Unit (VVCAU) which is utilized to abstain from smashing amongst vehicles. What's more, the other is Black Box (BB) which records the pertinent insights about the vehicle. For example, Distance from snag, Speed of vehicle, Brake status, Alcohol content, Accident Direction and excursion Time and Date. Presently a days' few works are done with respect to pre-crash identification & shirking framework from obstruction. Be that as it may, we are missing to needs crash with human or creatures contrasted with hindrances. What might be the circumstance when a framework needs to pick amidst human and impediments? In the event that the framework is not ready to identify human, then there is probability to hit people rather than impediments. To determine this issue in this work we have need people or creatures first. Auto will maintain a strategic distance from human or creature first then if conceivable it will attempt to keep away from obstructions moreover.

---

**Rajiv Gandhi Institute of Technology, Bangalore.**

---

**Title:** 3D Viewing of Snowman

**Tools:** Microsoft Visual C++

**Language:** C++ with OpenGL Package

**Description:** Project to show necessary steps and routines to the proper codification and simulation of a Land Of Snowman in 3D. With a simulation model, we can have a closer vision of the object of study in contrast with reality, what makes us capable of foreseeing how a determined object will look like and how it will behave after its proper construction in physical world.

---

**ACADEMIA**

---

**2013-2015 Master of Technology (Computer Science)** - Visvesvaraya Technological University – Rajarajeswari College of Engineering, Bangalore - Secured 67% (aggregate of 1<sup>st</sup> semester, 2<sup>nd</sup> semester and 3<sup>rd</sup> semester).

**2009-2013 Bachelor of Engineering (Computer Science)** - Visvesvaraya Technological University - Rajiv Gandhi Institute of Technology, Bangalore, India - Secured 62%.

**2009 Pre-University Course – (Physics, Chemistry, Mathematics, Computer Science)** - Karnataka PU Board S. Cadambi Independent PU College, Bangalore, India - Secured 73%.

**2007 ICSE - Class 10<sup>th</sup>** - S. Cadambi Vidya Kendra, Bangalore - Secured 71%.

## **PAPERS PUBLISHED IN JOURNALS AND PRESENTED IN CONFERENCE**

---

- [1] Dimpu Sagar N, Neelu L, **"IMPLEMENTATION OF DRIVER DROWSINESS DETECTION AND ACCIDENT AVOIDANCE IN VEHICLES – A REVIEW"**, published in International Journal of Innovative Research in Computer and Communication Engineering, Volume - 3, Issue - 4, April 2015.
- [2] Dimpu Sagar N, Neelu L, **"IMPLEMENTATION OF DRIVER DROWSINESS DETECTION AND ACCIDENT AVOIDANCE IN VEHICLES – A REVIEW"**, presented in International Conference on Emerging Trends in Engineering and Technology, April 2015.
- [3] Dimpu Sagar N, Dr.Usha Sakthivel, **"CREATING A MODEL FOR PREVENTING SEMANTIC EMAIL VIOLATIONS"**, published in International Journal of Combined Research & Development (IJCRD) eISSN: 2321-225X; pISSN: 2321-2241 Volume: 4; Issue: 3; April -2015.
- [4] Dimpu Sagar N, Mohan Kumar G, AAKANKSH S G, **"CDINS: COOPERATIVE PROVABLE DATA POSSESSION FOR INTEGRITY VERIFICATION IN NETWORK-CODING-BASED MULTI CLOUD STORAGE SYSTEM"**, published in International Journal Of Innovative Research & Development (IJIRD); JUNE -2015.

## **ACHEIVEMENTS & BEYOND CURRICULUM**

---

- ☞ Cleared the first 2 rounds of the Code-O-Pedia event conducted by the Programmers' Club of RGIT in the year 2010.
- ☞ Secured 2nd place in the event "Dot Exe" (Coding Event) conducted by RGIT Tech Fest – Agam 2011.
- ☞ For the Designs created entire year I was awarded Best District Committee Member for the year 2012-2013 from Rotaract District Council 3190.
- ☞ Held the post of Chief Coordinator in the Programmers' Club, a college club (2011 & 2012).
- ☞ Had been the Active Organizer of RGIT Tech Fest - Agam2012.
- ☞ Held the post of Designer in the Rotaract District 3190 (2012-2013).
- ☞ Held the post of Editor in Rotaract Club of RGIT (2012-2013).
- ☞ Held the post of Editor in Rotaract Club of Yelahanka (2014-2015).
- ☞ Held the post of Joint- Director PR-Designing in Rotaract District 3190 (2014-2015).
- ☞ Designed the website for Manavi Landscape ([www.manavilandscape.com](http://www.manavilandscape.com)).
- ☞ Designed the website for Pranavaa Essential Oils ([www.pranavaaoils.com](http://www.pranavaaoils.com)).
- ☞ Handling the position of Director Editor in Rotaract District Council 3190 (2015-2016).
- ☞ Participated in SDN Hackfest Conducted by ONF Foundation and ASM Technologies (2016).
- ☞ HSSE focal for Shell ARCC Tax and non-Finance Team.
- ☞ Champions on EPF Primer League 2019 Cricket Tournament in Shell.

## **PERSONAL DOSSIER**

---

Date of Birth : 10th August 1991

Passport Number : M4064230 with B1/B2 USA VISA valid till 2027.