Dheeraj Guglani
Associate L2 Data engineering
Publicis sapient
Gurgaon
Email:dhruv1954guglani@gmail.com
Mobile: +91 7404310273
SEP 2016 to till date

Summary of Experience and Qualifications

- **Primarily skilled in Hive, Sqoop, Impala, Python, PySpark, AWS GLUE, Athena with total experience near to 4 years.**

- **Worked on Technologies like IAM ,AWS EMR, Redis, Athena.**

- **AWS Certified Cloud Practitioner.**

- **Worked on creation of end to end pipelines using AWS EMR and Pyspark for data processing.**

- **Worked on Python, Pandas, NumPy and created an automation script that will compare the data with File to HDFS and will highlight the differences in the excel format.**

- **Successfully completed Introduction to Python, Intermediate Python for Data Science, Python Data Science Toolbox1 ,Python Data Science Toolbox2 certifications from DATACAMP.**

- **Successfully completed the 30 Days Bigdata training conducted by Deloitte.**

- **Successfully Completed CCA159(Cloudera Certified Data Analyst using Sqoop,Impala,Hive) on Udemy.**

- **Developed a Blitz Automation system using python that will automatically send the mails to the people and will automatically track the responses. Tool is currently being used by the HR team.**

- **Effective Team player with good communication skills, problem solving skills, analytical and computing abilities.**

- **Working knowledge of tools like Jira, Confluence, ControlM.**

Baxter
Role:Pyspark Developer
Duration: June 2019 to Present
Role:
- **Working on creation of end to end pipelines on Amazon EMR using Pyspark.**
- **Created Complex minus queries using Athena to test end to end pipeleines.**
- **Worked on build and testing of various objects and moved them to higher environments.**
- **Worked on migration of jobs in Pyspark that were already build in IBM DataStage.**
- **Migrated the data from legacy sources to the S3 buckets and processed them using Pyspark.**
- **Created the tables using Crawlers(created using AWS GLUE) and used them in Athena for SIT validations.**
- **Awarded Applause award in the project for showing commitment to the tasks and completed them before time.**

Exelon
Role: PySpark Developer
Duration: Aug 2018 to June 2019
Role:

- **As Part of Dev team took ownership of complete user stories and ensured to take them to the logical end.**
- **Worked on a tech stack including Hive,Sqoop, Spark,ODI and did the orchestration using UC4.**
- **USED UC4 to move the data among the different layers from Foundation to Consumption ,from Consumption to input/output layer.**
- **Used Hive,Sqoop to carry out the Business Transformations in F2C layer and moved the data from Foundation to Consumption layer.**
- **Used Pyspark to create end to end pipelines and moved the code to higher environments.**
- **Worked on F2IO(Grid Dev) where spark framework will convert the ODI generated parquet into JSON.**
- **Carried out the smoke testing in INT,UAT,PROD for the code developed and made sure the code is running fine for all the business scenarios.**
- **Awarded Spot award in the project for carrying out the F2IO work.**

  .

HPE(HP Enterprises)
Role(Automation Developer/Tester)
Duation: April 2018 to August 2018
Role:

- **Created and modified Python Scripts to check the DDL validation is happening correctly.**
- **Created hive queries to check the data coming from different sources is populated correctly in the Hive tables.**
- **Created and modified Python Scripts to check the Data validation is happening correctly.**
- **Created Test cases for the different objects and raised defects for the same.**
- **Automated a lot of functionality to smoothen the testing process.**
- **Imported the data from Oracle to Hive using Sqoop.**
- **Used Pyspark to carry out the testing for various objects and raised various defects in the SIT phase.**

Enhanced Personal HealthCare (EPHC)
Industry: Healthcare Insurance
Client description: Leading US Healthcare Insurance Provider
Duration: Dec 2017 to April 2018.
Project Description: An estimated one third of the total $2.7 Trillion spent on US Healthcare is driven by inappropriate care and administrative overheads. The market and increasing regulatory pressure demanded a new approach to focus on reducing the overall cost and to improve the quality of care. In order to achieve this, Deloitte implemented the Enhanced Personal Health Care Program (EPHC), which would - Transition from Fragmented to Integrated and Coordinated care by assigning a dedicated Physician to manage and coordinate the comprehensive care of each Member, shift from volume based to value based payments by increasing the liability of Providers towards overall care and incentivizing/penalizing them, based on their performance, and improve the quality of care by facilitating the Providers/ACOs with tools, resources and relevant information, around the needs of individual Members. The project aims at reducing 1% claims made each year translating to $100 million in savings.

Role:

- Understanding the Business and the Technical requirements to come up with Test plan, test ideas and test cases.
- Interacting with the clients, address issues and concerns, and provide solutions.
- Interacting with Client's Business team, Technical team and Development team to work on effort estimation, project planning and execution.
- Creation of automation queries within the project for each functionality.
- Carried out the testing for various objects based on the test cases and unleashed the defects during SIT phase.
- Presented Demos to the Product owner about the functionality implemented and how the user story
  Will help the business.

Accomplishments:
- Awarded SPOT award for showcasing commitment towards client deliverables.

BLITZ Automation Tool: Firm Initiative

Manager: Sudhakar Dhulipudi
Technology Used: Python (Pandas, Numpy, Win32Com)
Firm Initiative: The purpose of the initiative was to send automatic emails to the practitioners according to their skills so that they can take the interviews. Mails were send using win32com module of python with a response tracker and then the mails are read from the outlook of the HR to see how many people have responded and on the basis of which the final confirmation is sent to the practitioners. It made the recruitment process efficient and easy.
Role:
- **Responsible for creating the UI using Tkinter.**
- **Responsible for implementing the overall flow (Connecting the UI to the functionality).**
- **Responsible for reading the responses from outlook through win32Com module.**
- **Responsible for sending automatic emails to the practitioner to confirm their availability.**

.

Development Tools/Languages: `

- Hive   proficiency -Advanced
- Python   proficiency-Intermediate
- Teradata   proficiency-intermediate
- SQOOP   proficiency-intermediate
- Pandas   proficiency-intermediate
- Numpy   proficiency-intermediate
- Pyspark   proficiency-intermediate
- AWS   proficiency-Intermediate

.

Education:
BACHELORS OF COMPUTER APPLICATIONS
D.A.V Centenary College,Faridabad