

# DEIVAPRIYA S

Associate Consultant



Chennai, Tamil Nadu



[deivapriyasivasubramanian@gmail.com](mailto:deivapriyasivasubramanian@gmail.com)



+91 7358338079



<https://www.linkedin.com/in/deivapriya-sivasubramanian/>

## SUMMARY

- Associate Consultant with **5+ years** of experience in IT industry
- Proficient in **Python and Machine Learning**
- Worked on technologies like **Python, PySpark, Azure**
- Data analyst with hands on experience in analyzing the data using machine learning
- Worked on image processing techniques using Machine learning algorithm like neural network, SVM and experience in deep learning algorithms.



## SKILLS

- ML Algorithms : Linear regression, Logistic Regression, Decision tree, Naïve Bayes, Neural Network, Support Vector Machine
- Deep Learning Algorithm : Convolutional Neural Network
- Programming Languages : Python
- Frameworks : Flask, Pandas, Keras, Tensorflow, Numpy, Scikit-Learn, Pytest
- Data Processing : PySpark, Kafka
- DevOps Tools : GitHub, Jenkins, Databricks, Airflow
- Database : Postgres, Oracle, SQL, Hive, AWS S3
- AWS SDK : Boto3, Psycopg2



## PROFESSIONAL EXPERIENCE

May 2021 – **Associate Consultant**  
Present Infosys Pvt Ltd, Chennai

### 1. Log Analytics Solution:

- Developed an AI/ML driven Log analytics solutions (internal solution) to identify error patterns from huge volume of logs, detect error logs real time based on configured patterns and classify them for further processing
- Tech stack used – ELK Stack, **Hive, Kafka, Spark, MySQL**, Jenkins, LIVY, ML based models for identifying error patterns

## 2. Azure Model Development and Deployment Pipeline:

- Built and deployed various classification models using frameworks – **Sci-kit learn, Tensorflow, Keras and Pytorch**
- Used **Azure – Databricks** workspaces for developing and deploying the model.
- **Azure MLFlow** was utilized for Experiment tracking, artifact logging, registering model and model serving.
- Worked on deploying models using **Seldon** framework and in **AzureML** (real time/batch) based on custom scoring script

## 3. Model Management SDK/API:

- Developed a custom Python SDK which is built on top of open source MLflow to provides additional governance controls like signature validation on log Model and tags validation on the registered model.
- A custom python API using Flask for validating parameters that are given as input to SDK.

Jan 2017 – **Analyst**  
Apr 2021 Cognizant Technology Solutions, Chennai

## 1. Automated Ticket Classification using RNN networks and LSTM architecture

- **Description:** Optimized the help desk solution by integrating term-based prediction based on recurrent architecture by utilizing long short-term memory
- **Technology Used:** Python

## 2. Audio Generation using Generative Adversarial Networks:

- **Description:** Build an intelligent model which generates audio of different frequency by training a generative network (Convolutional neural network) which takes random noise vector as input and simultaneously trained discriminator network (Convolutional neural network) by feeding real audio as input which classifies the generated audio from the real. The classified results were in turn used to tune hyper parameters of generator to generate audios.
- **Technology Used:** Python

## 3. Intelligent Handwritten Character Recognition:

- **Description:** Presented a proof of concept to predict the handwritten characters (both isolated and cursive) from various forms and marked checkbox for Multiple-Choice Questions in the OMR sheet.
  - a. Neural network is the algorithm used and MNIST dataset was used to train the network. This helps in reducing the manual work of extracting the information from general forms and to analyze them.

- b. Support Vector Machine Algorithm was used for implementing the prediction of handwritten characters and clustering for predicting the marked checkbox. This helps in automating the OMR Sheet evaluation.
- **Technology used:** Python, Matlab

#### 4. **Webpage Automation:**

- Automated the process of fetching the data from a website and update it in a database
- Utilized redshift and oracle to retrieve and store the data
- Technology used: Python, selenium, Oracle, Redshift

#### 5. **Data Migration:**

- Migrated data used in generating report from SAS system to cloud (AWS S3)
- Perceive SAS scripts and recreate the same in SparkSQL/Pyspark technology
- Generated report based on weekly, monthly and quarterly basis based on business requirement
- Technology Used: PySpark, SAS, Airflow, Databricks, AWS S3, Jenkins, GIT

#### 6. **JSON Automation:**

- Developed a system for Automated JSON file for updating the data ingestion meta data information from various sources. Analysis of business requirement and automate the generation of JSON file thereby reduces the manual efforts
- Technology Used: Python, Airflow, AWS S3

#### 7. **Maintenance Dag for Airflow Scheduler:**

- **Description:** Developed Maintenance workflow dag to
  - a. periodically kill tasks that are running in the background that do not correspond to running task in the database
  - b. periodically clean out task logs to avoid getting too big logs
  - c. save logs only for specified period and remove all old logs which consumes more space
- Technology Used: Python



## **EDUCATION**

**M.Tech Data Science and Engineering**  
BITS, Pillani

2019 – 2021

CGPA: 7.6

**B.E Electronics & Communication Engineering**  
PSNA College of Engineering & Technology, Dindigul

2012 – 2016

CGPA: 8.3



## **ACCOMPLISHMENTS**

- Participated in Hackathon to convert Unstructured document (PDF) to Structured document (Excel) using python script which could help any downstream systems that consumes Excel for analysis
- Awarded as Tech Wizard 2022 in Infosys