

# MOUNA P

## Data Scientist

E: [msrisai.p@gmail.com](mailto:msrisai.p@gmail.com) | M: +44 7468630610 | L: [LinkedIn](#) | A: London, UK

### Summary

---

A Machine Learning Engineer with **4+years** of experience in Big Data Analytics, Data manipulation, Data Mining and implementing Machine Learning models to solve various business problems.

- Excellent understanding of **SDLC, Agile and Scrum** development methodology.
- Experience in Machine Learning with large data sets of **Structured and Unstructured data**, Data Acquisition, Data Validation, Predictive modeling, Data Visualization.
- Proficiency in data warehousing inclusive of dimensional modeling concepts and in scripting languages like **Python, Scala, and JavaScript**.
- Experienced in machine learning libraries and packages, such as **TensorFlow, Keras, scikit-learn, OpenCV, Flask, NumPy, Pandas**.
- Strong understanding of **OOP** Concepts and good knowledge of common Design Patterns.
- Good experience with deep learning frameworks like **TensorFlow and PyTorch**, focusing on **CNNs, RNNs, GANs**, and Transformer architectures.
- Experience dealing with large data sets (**structured, semi-structured and unstructured**), cleaning the data and organizing the data for analyses.
- Experience with time series data, and perform various statistical and (**AI**) **artificial intelligence, (ML) machine learning algorithms** such as LMS, regression, filtering, correlation, neural network.
- Gained experience to develop API libraries and coded business logic using **C#, XML** and designed web pages using **.NET** framework, **C#, Python, Django, HTML, AJAX**.
- Experience level knowledge on **python libraries** like **NumPy, Scipy, Pandas, Seaborn** used for Data Analysis.
- Have experience with various visualization tools like **Tableau, R, Plotly, Power BI** to understand the insights.
- Hands on experience in Data Analytics Services such as **Athena, Glue, Quick Sight**.
- Experienced in implementing **Machine Learning algorithms** to develop predictive models.
- Working experience in **SQL Server, MySQL, Spark SQL, PostgreSQL, MongoDB, Cassandra, DynamoDB**.
- Gained experience on working and approaches in **Natural Language Processing(NLP)**.
- Good experience on **Deep Neural Networks, Convolutional Neural Networks(CNN), Recurrent Neural Networks(RNN)**.
- Strong experience of deep learning frameworks such as **TensorFlow or PyTorch** for building and training neural networks.
- Skilled in data visualization tools like **Power BI, Tableau and Matplotlib** to communicate findings and present complex data in a clear and compelling manner.
- Gained Familiarity with **natural language processing (NLP) techniques** and tools for text mining and sentiment analysis.
- Experienced in using distributed computing architectures such as **AWS** products (**e.g. EC2, Redshift, and EMR, Elasticsearch**), **Hadoop, Python, Spark** and effective use of **map-reduce, SQL** and **Cassandra** to solve big data type problems.
- Performed Information Extraction using **NLP** algorithms coupled with **Deep learning (ANN and CNN, RNN, LSTM, Encoders)**, **Keras and TensorFlow**.
- Extensive experience on other machine learning libraries such as **Seaborn, Scikit-learn, SciPy** for machine learning, familiar working with **NLTK, spaCy** for **Natural Language Processing**.

- Implemented **CI/CD with GitHub** Actions to push code updates to development, staging, and production Docker images, deploying them to the **Amazon EC2 registry**.
- Experience in **AWS Lambda, EC2, EMR, AWS Cloud Watch, Sage Maker, Redshift, AWS Step functions, Secret Manager and S3 Services**.
- Maintenance and monitoring of **Docker** in a cloud-based service during production and Set up a system for dynamically adding and removing web services from a server using **Docker**.
- Attend the project review meetings, daily/weekly status meetings and walkthroughs with the project manager in **JIRA**.
- Experienced the full software life cycle in **SDLC, Agile, Devops and Scrum** methodologies including creating requirements, test plans.

## Skills

---

**Natural Language Processing** - NLTK, Apache Spark, Scikit-learn

**Data Sources** - SQL Server, Excel

**Big Data Ecosystems** - Hadoop, MapReduce, HDFS, HBase, Hive, Spark

**Data Visualization** - PowerBI, TABLEAU, Matplotlib, Plotly, Seaborn, Matlab

**Deep Learning** - CNN, RNN

**Cloud Technologies** - AWS, DevOps

**Programming Languages** - Python, SQL, R, Javascript

**Scripting Languages** - XML, Python

**Version control** - Git, SVN (subversion), Github

**Operating System** - Linux, Windows, Unix

**Methodologies** - Agile, Scrum, Waterfall

**Bug tracking Tool** - Jira

## Work Experience

---

**Company:** Mind Tree

**Client:** Bank of America

**Role:** Data Scientist/Machine Learning Engineer

**Duration:** Aug 2022 to till date

**Description:** Designing, architecting, and implementing scalable and secure software solutions using Python, SQL or NoSQL, AWS and related technologies with the existing banking platform Implemented robust data processing and analysis solutions using Python libraries and frameworks such as Pandas, NumPy, or SciPy.

### Responsibilities:

- Have Knowledge of the Software Development Life Cycle (**SDLC**), **Agile and Waterfall** Methodologies
- Involved in different phases of **data acquisition, data collection, data cleaning**, model development, model validation, and visualization to deliver solutions.
- Designed the **(AI) artificial intelligence, (ML) Machine learning** data pipeline for regular monitoring and performance evaluation of the deployed **ML** models.
- Developed machine learning models using recurrent neural networks – **LSTM** for time series, predictive analytics.
- Involved in various Object-Oriented Programming (**OOPS**) concepts and implementing the same for Design and Architecture of the application.

- Written python code for exploratory data analysis using **Scikit-learn machine learning python packages- NumPy, Pandas, Matplotlib, Seaborn, statsmodels, pandas profiling**.
- Developed machine learning models using **Google TensorFlow keras API** Convolution neural networks for Classification problems, fine-tuned the model performance by adjusting the epochs, bath size, Adam optimizer.
- Used **Python (NumPy, SciPy, pandas, Scikit-learn, seaborn) and R** to develop a variety of models and algorithms for analytic purposes.
- Worked on Natural Language Processing with **NLTK** module of python and developed **NLP** models for sentiment analysis
- Designed and implemented open source **AI frameworks - Pytorch, TensorFlow, Scikit-learn, Apache Open Source (Kafka, Storm, Spark) for NLP and Machine Learning Algorithms**.
- Performed data imputation using **Scikit-learn** package in **Python**. Programming in one or more of **Python, Java, .Net, JavaScript, C++**.
- Build Back- end support for Application from ground up using **Python, Shell scripts & Perl**. Design and implement custom **Bash scripts** for controlling the data flow from the **PostgreSQL database**.
- Used **AWS Cloud Watch** for performing Monitoring, customized metrics and file logging and successfully completed the task in using **AWS Lambda** to manage the servers and run the code in the **AWS**.
- Implemented **CI/CD** processes using **Jenkins** ensuring efficient and reliable software delivery
- Successfully migrated the **Django** database from **SQLite to MySQL to PostgreSQL** with complete data integrity.
- **Deep learning (LSTM** recurrent neural network (**RNN**)) package was used for sequence to sequence (Seq2Seq) **LSTM** model.
- Analyzing various logs that have been generating and predicting/forecasting the next occurrence of events with various **Python libraries**.
- Involved with creating user stories and issues, tracking team status, scheduling user tasks in a sprint using **JIRA**.

**Environment:** Python, AI, ML, LSTM, OOPS, Scikit-Learn, Numpy, Pandas, Matplotlib, Seaborn, stat models, API, Tensorflow, Keras, NLTK, NLP, Pytorch, Kafka, Java, Javascript, C++, Shell scripts, Perl, Bash, CI/CD, Jenkins, AWS Lambda, Django, SQLite, My SLQ, PostgreSQL, RNN, Python Libraries, Jira.

**Company:** IBN Technologies

**Client:** Axis Bank

**Role:** Python Developer

**Duration:** Sep 2020 to Aug 2022

**Description:** Contributed to the implementation of business logic, facilitated seamless data exchange, performed XML processing, and created engaging graphics using Python and Django. Implemented RESTful APIs using frameworks like Django Rest Framework and Flask-RESTful, enabling seamless data exchange between the frontend and backend.

#### **Responsibilities:**

- Worked on Data mining with large datasets of Structured and Unstructured data, Data Acquisition, Data Validation, Predictive modeling, Data Visualization.
- Developed web applications using **Python, Flask, and the MVC** architecture ensuring seamless functionality and optimal user experience.
- Used **Python (NumPy, SciPy, pandas, Scikit-learn, seaborn)** and R to develop a variety of models and algorithms for analytic purposes.

- Automated the **ML** model building process by building Data Pipelines further integrating it with Data cleaning process
- Integrated third-party libraries and modules, such as **Pandas and NumPy** for data manipulation and **TensorFlow** for machine learning, to enhance website functionality and data processing capabilities.
- Developed entire frontend and backend modules using **Python on Django Web Framework**
- Worked extensively with **Bootstrap, JavaScript, and JQuery** to optimize the user experience. Used **Python Library BeautifulSoup** for web scraping to extract data for building graphs.
- Developed **UI using CSS3, HTML5, JavaScript, AngularJS, JQuery and JSON.**
- Worked with **Amazon Web Services (AWS) using EC2** for hosting and Elastic map reduce (**EMR**) for data processing with **S3** as storage mechanism.
- Experience in working with databases, such as **MySQL or PostgreSQL**, and knowledge of **SQL** queries and ORM (Object-Relational Mapping) tools.
- Developed SQL Queries, Stored Procedures, and Triggers Using Oracle, **SQL and PL/SQL.**
- Worked on data preprocessing, developing different statistical machine learning models and data mining solutions to various businesses, generating data visualizations using **Python, R, Tableau, Microsoft Power BI, version control with GIT.**
- Integrated **SOAP and RESTful APIs** to facilitate efficient communication between different system components, ensuring data consistency and reliability.
- Implemented a **CI/CD pipeline** with **Docker, Jenkins and GitHub** by virtualizing the servers using **Docker** for the Dev and Test environments by achieving needs through configuring automation using Containerization.
- Use **Jira** for project management teams such as report,& analysis, workflow customization, issue /task management, project customization and help to team of all types manage work.
- Implemented **Agile** Methodology for building an internal application.

**Environment:** Python, Git, Github, CSS# HTML, Javascript, Angular JS, Query, Json, MySQL, PostgreSQL, PL/SQL, R, Tableau, Power BI, Soap, Restful API, CI/CD, Jenkins, Github, Flask, MVC, Numpy, Scipy, Pandas, Scikit, Learn, Seaborn, ML, AWS, DevOps, Agile.