

Saipranay Masadi

Data Scientist

Herndon, VA | 571-436-5367 | saipranay@myjobtech.com | [LinkedIn](#) | [GitHub](#)

SUMMARY

- Around 5 years of experience as a data scientist and expertise in programming languages, specializing in data manipulation and analysis.
- Proficient in utilizing a wide range of tools and frameworks, including Scikit-Learn, Streamlit, TensorFlow, Keras, NLTK, OpenCV, and Spark for machine learning, natural language processing, computer vision, and scalable data processing tasks.
- Experience in hypothesis testing, regression analysis, clustering, classification techniques, decision trees, support vector machines and time-series analysis, facilitating accurate predictive modelling and inference.
- Understanding statistical concepts and methodologies and practical experience applying them to real-world data science projects.
- Continuously stays updated with the latest advancements in data science, demonstrating a commitment to professional growth and the pursuit of excellence.

SKILLS

Methodologies: SDLC, Agile, Waterfall

Language: Python, R, SQL, SAS, MATLAB

IDEs: Visual Studio Code, PyCharm, Jupyter Notebook

Packages: NumPy, Pandas, Matplotlib, SciPy, Scikit-learn, Seaborn, Pytorch, TensorFlow, Ggplot2, Plotly

Machine Learning Algorithm: Linear Regression, Logistic Regression, Decision Trees, Supervised Learning, Unsupervised Learning, Classification, SVM, Random Forests, Naive Bayes, KNN, K Means, CNN, Natural Language Processing (NLP)

Cloud Technologies: AWS, Azure, GCP, Databricks

Visualization Tools: Tableau, Power BI

Database: MS SQL Server, MySQL, SQLite

Other Tools: Git, MS Office, Google Analytics

EDUCATION

MS in Information Systems Technology | Wilmington University

Aug 2022 - Dec 2023

Bachelor of Engineering | Osmania University

Sep 2014 - May 2018

EXPERIENCE

Freddie Mac, USA | Data Scientist

Aug 2023 - Present

- Implemented and deployed various machine learning algorithms, including linear regression, logistic regression, decision trees, and K-means clustering to address supervised and unsupervised learning objectives.
- Conducted thorough data analyses using Python and Tableau to derive actionable insights, focusing on achieving accuracy in monitoring dependent variables and identifying key factors impacting match success rates.
- Managed EC2 instances, S3 storage, and RDS databases, ensuring optimal performance and scalability.
- Led the integration of data pipelines from web servers and Teradata systems using Sqoop in conjunction with Kafka and Spark Streaming API, facilitating seamless data ingestion and processing.

Hexaware Technologies, India | Data Scientist

Jan 2020 - Jul 2022

- Actively participated in Agile methodologies, contributing to various scrum ceremonies such as sprint planning, daily scrum, sprint review, and retrospective meetings.
- Applied data cleaning, feature scaling, and feature engineering techniques using pandas and NumPy in Python, optimizing data quality for model development.
- Developed statistical algorithms, including Multivariate Regression, Logistic Regression, PCA, Random Forests, Decision Trees, and Support Vector Machines using Python, R, and SQL to assess welfare dependency risks.
- Utilized Teradata SQL Workbench to craft custom SQL queries, extracting and preparing datasets tailored for Tableau dashboards.
- Leveraged deep learning frameworks for various data-driven tasks to build and fine-tune models, including deep neural networks, Convolutional Neural Networks (CNNs), and Recurrent Neural Networks (RNNs).

Streebo Inc, India | Data Scientist

Jun 2018 - Dec 2019

- Conducted exploratory data analysis using Python, R, SQL, and SAS to uncover patterns and trends in complex datasets.
- Developed and implemented supervised and unsupervised machine learning models using Scikit-learn, TensorFlow, and Keras to solve business problems.
- Leveraged Databricks for scalable data processing and analysis, efficiently handling large volumes of data.
- Designed and optimized relational databases using MS SQL Server, MySQL, and SQLite for efficient data storage and retrieval.
- Utilized Git for version control, managing codebase efficiently, and facilitating collaboration with team members.

CERTIFICATIONS

Machine Learning Specialization- Coursera Stanford University Online – by Andrew Ng

Machine Learning System Design - Educative

Generative AI with Large Language Models - DeepLearning.AI

MongoDB Python Developer Path

Learn Embeddings and Vector Databases- Scrimba