Mamesa El

+1(206) 403 0785 | Seattle, WA

Email: mamesa.el@berkeley.edu | LinkedIn: linkedin.com/in/mamesael | GitHub: https://github.com/Mamesa-El

Professional Profile

As a UC Berkeley Data Science Master's graduate, I have cultivated a solid foundation in programming, mathematics, and machine learning, closely aligning with the skill set of a Data Scientist. My tenure as a Data Scientist at Volt has sharpened my analytical abilities, particularly in tackling physics-centric applications. This blend of advanced academic training and practical experience in data-driven environments positions me to make substantial contributions as a Data Scientist, especially in dynamic and data-intensive projects.

Work Experience

Data Scientist   April 2023 – Present
Volt   *Tulsa, Oklahoma*

- Developed and managed a Scikit-Learn, NLP, and BERT-based spam detection system, achieving an 80% F1 score through extensive statistical analysis for improved accuracy.
- Developed SMS data pipelines, encompassing tasks such as data preprocessing, feature engineering, and model selection, to optimize model flow efficiency.
- Managed machine learning model lifecycles, including AWS SageMaker deployment, enhancing cloud deployment experience.

Data Analyst   Sep 2019- June 2021
Department of Physics and Astronomy, Western Washington University   *Bellingham, WA*

- Collaboratively analyzed spectroscopic data from 12 galaxy clusters using Python and Astropy to extract key properties.
- Implemented anomaly detection models for galaxy mass, luminosity, and density, strategically targeting extreme luminosity variations.
- Developed a script to efficiently analyze and visualize over 200,000 galaxies, showcasing practical ML solutions development.

Projects

Flight Delay Prediction:

- Improved United Airlines' flight delay prediction model by concentrating on feature engineering and model evaluation with XGBoost, achieving an impressive 77.2% F1 score.
- Employed statistical analysis techniques, including bootstrap sampling, to derive generalizable insights from data samples.
- Processed large-scale weather datasets in Parquet format, significantly enhancing the accuracy of predictions.

Skills

Programming: Python, SQL, Bash (Linux Shell Scripting)
Frameworks/Packages: TensorFlow, Pytorch, LangChain, SK-Learn, Machine Learning, Pandas, NumPy, Matplotlib
Platforms and Tools: Databricks, AWS SageMaker, Spark, Hadoop
Containerization/Orchestration: Docker, Microsoft Azure Kubernetes
Version Control: Git

Education

University of California, Berkeley – Master of Information and Data Science   December 2023
Relevant Courses: Machine Learning at Scale, Machine Learning System Engineering, Applied Machine Learning, Statistics for Data Science, Fundamental of Data Engineering

Western Washington University – Bachelor of Science in Physics   June 2021