YASMIN BADRI

Data Scientist | (707) 386-3756 | ywbadri@gmail.com

SUMMARY

Data Scientist with 3+ years of experience in python development, data cleaning and validation, and testing. Certified in Data Science with expertise in Python, SQL, Data Visualization, Data Modeling, and AWS. Strong background in building ETL workflows and predictive and classification models.

SKILLS

Python, SQL, PySpark, AWS S3, Docker, Kubernetes, CI/CD

ETL/ELT, Data Pipelines, API Integration, Data Validation

Scikit-learn, TensorFlow, Keras, NLP, Computer Vision

Looker, Matplotlib, Seaborn

WORK EXPERIENCE

DATA QUALITY ENGINEER

CoinMetrics (Remote) | May 2022 - Present

- Developed Python web applications for blockchain data validation
- Automated data quality checks using PySpark and AWS S3
- Reduced client-reported issues by 70% through validation tools
- Implemented backend features with Docker and Kubernetes

DATA SCIENCE IMMERSIVE STUDENT

General Assembly (Remote) | Aug 2021 - Nov 2021

- Trained CNN image classifiers using TensorFlow/Keras
- Developed regression models for health outcome prediction
- Built NLP models for Reddit post classification

QUALITY ASSURANCE ANALYST

Le Tote (Shenzhen/Stockton) | Sep 2016 - Jul 2020

- Created Looker dashboards for operational performance tracking
- Automated reporting processes, saving 40% manual effort

PROJECTS

FOOD IMAGE CLASSIFICATION

- Developed CNN models for food image recognition

HOUSING PRICE PREDICTION

- Built regression models for Iowa housing price forecasting

EDUCATION

DATA SCIENCE IMMERSIVE CERTIFICATE

General Assembly | 2021

DATA SCIENCE IN PYTHON

Dataquest.io | 2020-2021