

Eric (Ross) Fu

713-540-4528 | ericrossfu@yahoo.com | Houston – Open to Relocation (Los Angeles Miami Austin preferred) | [LinkedIn](#)

SUMMARY

Proactive and collaborative data scientist with 7.5 years of experience leveraging machine learning, statistical modeling, and data-driven strategies to solve complex problems. Passionate about uncovering actionable insights and delivering high-impact solutions. Adept at end-to-end model development, from data preprocessing to deployment, with a strong foundation in Python, SQL, and cloud-based ML pipelines.

EDUCATION

Southern Methodist University

Master of Science in Data Science

Dallas, Texas

December 2020

Texas A&M University

Bachelor of Science in Biology

College Station, Texas

December 2015

EXPERIENCE

Data Scientist

2021-Present

Avance Biosciences Inc.

- Hired as the first engineer bridging data science, data engineering, and full-stack development. Spearheaded innovative solutions that transformed core practices, creating a track record of high-impact improvements.
- Trained a cutting-edge AI model on proprietary company data, harnessing advanced LLM capabilities to deeply understand and extract insights from company materials.
- Deployed a state-of-the-art AI model with a seamless web interface and robust API integration, enabling intuitive user interactions and efficient, precise data retrieval.
- Deployed a suite of advanced machine learning models via AWS Sage Maker, driving automated predictions that enhanced decision-making efficiency by up to 50%, simultaneously reducing errors and optimizing workflows.
- Engineered a sophisticated data pipeline leveraging advanced computer vision techniques, boosting data intake capabilities by 10x.
- Took responsibility for the organization's entire data management, engineering and deploying Kafka, Spark, and AWS pipelines for ETL and backup, ensuring data integrity and governance across 70+ sources and file types.
- Quadrupled data capture in the company database, developing critical data resources that enhanced data governance and streamlined operational coordination across teams.
- Developed multiple Python-based applications, incorporating APIs, front-end GUIs, and back-end functionality to replace core processes and address key problem areas.
- Empowered users to conduct a core process with a 600-fold efficiency boost, minimizing human error, delivering real-time actionable insights and enhancing operational coordination.
- Catalyzed a company cultural shift toward automation, scalable practices and data governance.
- Technologies Used: Python, MySQL, AWS, FastAPI, Machine Learning, AI, Docker, Kafka, Spark

SKILLS

- Programming Languages: Python, SQL, R
- Machine Learning: TensorFlow, PyTorch, Keras, HuggingFace, AI, LLM, NLP, Neural Network
- AWS: Kinesis, Glue, EMR, S3, Lambda, Sagemaker, Redshift etc.
- Azure: DataFactory, Databricks
- Apache: Kafka, Airflow, Spark
- API: FastAPI, Streamlit
- Other: GitHub, Docker, Statistics