

Jose Mancera

[linkedin.com/in/josemancera](https://www.linkedin.com/in/josemancera) | 979-479-1776 | mancerajose.com | mancerajose11@gmail.com

EDUCATION

Baylor University, School of Engineering & Computer Science – Waco, TX
Bachelor of Science in Data Science

Double Major: Statistics

Minor: Economics

May 2025

TECHNICAL SKILLS

Deep Learning & Programming: Python, C++, CUDA, Java, PyTorch, TensorFlow, Keras, Generative Models, SQL, R, TypeScript

Data Science: Pandas, NumPy, Scikit-learn, Statistical Modeling, A/B Testing, Apache Spark

Cloud, DevOps, & Tools: Docker, Kubernetes, Apache Airflow, Grafana, AWS (Lambda, SageMaker, S3, ECS), Git, Jupiter Notebooks

EMPLOYEMENT

Amazon - Software Development Engineer, Intern – Santa Clara, CA

Jun 2024 – Aug 2024

- Developed and supported SageMaker LLM inference and Generative AI SDKs, resulting in a 43% reduction in inference time.
- Enhanced UX and support for Quantization Algorithms, streamlining quantizing models for various backends in LMI using CUDA.
- Developed a TensorRT-Model-Optimizer pipeline with quantization, sparsity, distillation, and pruning, increasing features by 180%.
- Tested algorithms like AWQ, squeezeLLM, smoothquant, gptq, fp8, converting quantized artifacts for Large Mode Inference (LMI).
- Performed benchmarking tests on quantized models, improving resource utilization by 40% and cutting operational costs by 20%.

SpaceX - Software Engineer, Intern - McGregor, TX

Jan 2024 – Apr 2024

- Led the design and implementation of a microservices-based platform, ensuring scalability and reliability with Kubernetes.
- Implemented an ETL pipeline with Apache Spark, NiFi, and Airflow, improving SQL Server performance and data handling by 57%.
- Developed and designed an E-Invoice app to test and pull data from 3rd party vendor APIs, integrating data processing.
- Utilized Grafana to monitor performance metrics of microservices and ETL pipelines, achieving a 2x performance improvement.

Amazon - Software Development Engineer, Intern - Seattle, WA

May 2023 – Aug 2023

- Migrated internal diagnostic tool to serverless architecture using AWS Lambda and ECS Fargate.
- Collaborated with cross-functional teams to enhance the tool's performance and scalability significantly by 63%.
- Assessed AWS services and integrated Amazon ECS for enhanced container orchestration.

United States Marine Corps – Work Center Supervisor, Sergeant - MCBH Kaneohe Bay, HI

Nov 2015 – Nov 2020

- Supervised and managed 35+ individuals and served as the final quality check on maintenance for all life support equipment.
- Trained personnel to complete 7 major maintenance qualifications, increasing division skillset by 25%.
- Composed and produced written and oral communications between 7 different flight organizations.
- Managed 463 life support equipment assets valued at over \$1.3 million.

PROJECTS

Generative AI Models for Image Generation

Jan 2024 – Apr 2024

- Developed and trained a GAN-based model using TensorFlow and Keras for high quality image generation, achieving 97% accuracy.
- Implemented a Transformer-based model with PyTorch for generating coherent and contextually relevant text, trained on big data.
- Evaluated GAN and Transformer models using Inception Score (IS), Fréchet Inception Distance (FID), BLEU, and ROUGE scores.
- Explored the application of the Transformer model for audio generation, synthesizing audio waveforms from text inputs.

Web3.0 NFT Marketplace with Generative AI

Aug 2023 – Nov 2023

- Developed an NFT Marketplace application using Next.js, Solidity, and Polygon blockchain, with integrated crypto wallet features.
- Integrated a GAN-based model with TensorFlow and Keras to enable users to generate unique, high-quality NFT artwork.
- Implemented a Transformer-based model in PyTorch for generating detailed textual descriptions and metadata for NFTs.

Portfolio Recommendation System

Dec 2022 – Feb 2023

- Developed a recommendation system for creating stock portfolios using machine learning algorithms, generating an alpha of 23%.
- Leveraged XGBoost to enhance predictive accuracy and incorporated indicators to construct diversified portfolios.
- Implemented a backtesting framework to simulate portfolio allocations and trading strategies, leading to 18% profit margins.
- Utilized TensorFlow and a Transformer framework to build and train deep learning models for improved prediction accuracy.

RELEVANT COURSEWORK

CS + DS: Algorithms and Data Structures, Machine Learning Theory, Analytics for Machine Learning, Statistical Machine Learning, Cloud Computing, Big Data Analytics, Database Design, Software Project Management, Data Visualization, Distributed Systems

Math + Stats: Multivariable & Vector Calculus, Linear Algebra, Differential Equations, Statistics & Probability, Regression Analysis, Bayesian Statistics, Applied Time Series, Probability Theory & Statistics, Statistical Inference & Analysis, Foundations of Higher Math

Economics: Game Theory, Energy Economics, Principles of Microeconomics, Principles of Macroeconomics