Cristian Hohbein

chohbein2@gmail.com | (714) 510-1314 | Brea, CA 92823 | Github: @chohbein

Education

University of Colorado, Boulder - *Master of Science in Data Science, Machine Learning*

September 2024 - Present

- **Relevant Coursework: Introduction to Deep Learning, Introduction to Machine Learning: Supervised Algorithms, Unsupervised Algorithms in Machine Learning**

University of California, Santa Barbara - *Bachelor of Arts in Data Science*

September 2022 - June 2024

- **Relevant Coursework: Advanced Statistical Models, Big Data Analytics, Statistical Machine Learning**

Frameworks, Libraries, & Platforms

- **PyTorch, TensorFlow, scikit-learn, NLTK, spaCy, YOLO**
- **AWS (EC2, RDS, Lambda, S3), Git, Docker**

Languages

- **Proficient: Python, SQL Competent: R, JavaScript, Bash**

Projects

- Cancer Detection CNN
  - Built a CNN in PyTorch + CUDA to detect breast-cancer in 220,000 histopathology images, achieving 0.98 ROC-AUC on a test set with only 2% false-negatives while training locally.
  - **Techniques:** Mixed-Precision (autocast / GradScaler), batch-norm & dropout, multi-worker dataloaders, image augmentations, early-stopping & learning rate experimenting, GPU batch-size tuning.
- News Aggregator
  - Built an AWS-based data-pipeline that scrapes and stores hundreds of news articles per day (EC2 ➔ S3 ➔ Lambda ➔ PostgreSQL RDS); applied NLP to cluster similar stories, extract keywords and generate multi-document topic summaries.
  - **Techniques:** Sentence-BERT similarity + DBSCAN, spaCy POS/NER, TF-IDF & PRIMERA summarization, NLTK keyword scoring, AWS EC2/S3/Lambda/RDS, PostgreSQL.
- Monet Painting GAN
  - Engineered and benchmarked two GANs for unpaired style transfer, culminating in a U-Net & WGAN-GP model that achieved a target MiFID score 3x faster than a baseline CycleGAN implementation.
  - **Techniques:** Wasserstein GAN with Gradient Penalty (WGAN-GP), asymmetric critic updates, CLIP-based perceptual loss (style and content), U-Net generator with CLIP-conditioned bottleneck, latent noise injection, spectral normalization, gradient clipping, and mixed precision (AMP).