

Sebastian Prestia

sprestia2009@gmail.com | (631) 278-0443

EDUCATION

University of California, Irvine – Masters of Data Science (on Fellowship)

Arizona State University – BS Data Science (Summa Cum Laude)

Graduated May 2026 at 16 • GPA: 4.0 Major, 3.99 Overall

Coursework: Computer Vision, Machine Learning, Linear Algebra, Probability, Multivariable Calculus, Statistical Computing

RESEARCH & PROJECTS

- **Sparse Symbolic Regression Research Project (In Progress):** A paid research project under ASU regrading Sparse Symbolic Regression, its limitations, and how it can be improved.
- **Financial Time-Series Pattern Recognition (In Progress):** A major project involving massive data collection into the 100s of gigabytes, tons of model experimentation, and pattern recognition to predict Stock Market movement.
- **Mole Detection:** Developed a Computer Vision model to classify moles as Benign or Cancerous on the ISIC Dataset. I experimented with multiple different visionary model architectures and training techniques, both my own and ones derived from research papers, and solved the class imbalance between Benign and Cancerous images to balance training. I further selected augmentations to prevent overfitting without losing the delicate details that distinguish Benign moles from Cancer.
- **JEPA vs Autoencoders:** Coded and Trained a Joint-Embedding Positional Architecture from scratch and compared the storage of information in its latent space to that of a Visionary Autoencoder's on CIFAR10 & 100.
- **JEPA Facial Recognition:** Coded and Trained a Joint-Embedding Positional Architecture model from scratch on Celebrity faces to form a latent space based on similar features.
- **Movie Box Office Prediction (Ridge Regression):** Built a regularized linear model using log transformed predictors; analyzed multicollinearity and model performance.
- **Syrian HDI War Impact Study (ANOVA + Regression):** Examined pre/post-war differences in HDI contributors with World Bank data.
- **NYC Taxi Analysis:** Used Statistical Techniques such as Clustering and KDE Heatmaps to identify best neighborhoods in NYC for taxis.
- **PM2.5 levels across US:** Analyzed PM2.5 levels across US and how factors such as Income and Urban Areas affect it.
- **CIFAR-10 / CIFAR-100 CNNs:** Designed deeper convolutional models; analyzed generalization limits without pretraining.

- **MNIST Digit Classifier (CNN):** Achieved 99% accuracy using convolutional architectures with dropout/batch norm.
 - **FashionMNIST Classifier (CNN):** Achieved 97% accuracy; studied generalization vs. MNIST.
 - **Custom Fish Dataset (CNN):** Built augmentation pipeline for small-data classification.
 - **NumPy Neural Network:** Implemented forward/backpropagation manually to understand gradient mechanics and optimization.
-

EXPERIENCE

Hydra Capital Group — Statistical Machine Learning Intern

- Developing a pattern-recognition AI model to identify recurring structural patterns in financial timeseries data, such as local U-shapes, breakouts, and reversal formations, with out-of-sample evaluation to assess generalization.
- Conducted structured backtesting on historical market data to assess model stability and refine pattern-based predictive performance.

WaxMelters — Machine Learning & AI Intern

- Developed AI chatbot tools and optimized website performance using statistical analysis techniques and data-driven evaluation.
-

TECHNICAL SKILLS

Programming: Python (advanced), PyTorch, NumPy, Pandas, Matplotlib, R, SQL

Mathematical Tools: Optimization methods, vector calculus, matrix computations

Machine Learning: CNNs, ANNs, supervised learning, regularization techniques, model validation, data preprocessing, augmentation

ACCOMPLISHMENTS

- Bronze Medalist — Pan-American Youth Chess Championship 2017
- Represented USA at World Cadet Championship 2017
- Candidate Master in Chess
- Chess Expert (2000+ rating); 2x National Champ, #1 in U.S., #2 World for age group (multiple years)
- Entered university full-time at 12
- Accepted into University of California, Irvine's Masters of Data Science program on Fellowship at 16
- Summa Cum Laude Graduate at 16