Johnathan Aguilar *Last updated in September 2025*

Education

California State University, Northridge *Sept 2023 – May 2026*
BS in Computer Science, Minor in Data Science
- GPA: 3.95/4.0

Moorpark College *May 2020 – May 2023*
Associate of Science for Transfer, Computer Science
- GPA: 4.0/4.0

Research Experience

Agentic AI Research Intern *Sun Valley, CA*
Waste Management & Autonomy Research Center For STEAHM *Sept 2025 – Present*
- Leading an undergraduate team in collaboration with a PhD student and a faculty advisor, coordinating research tasks and ensuring alignment with project goals.
- Developing an automated pipeline to transform handwritten reports into Neo4j Cypher code.
- Developing a knowledge base for AI by integrating Cypher queries from a Neo4j graph to make context-aware generations to support downstream digital twin simulations and help Managerial staff with intelligent decision-making.

Undergraduate Researcher *Northridge, CA*
California State University, Northridge - Department of Computer Science *Oct 2025 – Present*
- Conducting spatio-temporal analysis of Level 4 Sea Surface Temperature (SST) data from NASA PO.DAAC as part of a NASA JPL-sponsored research project.
- Using EOF/PCA with unsupervised clustering models to identify dominant spatial patterns and group regions with similar SST variability across time.
- Applying spatial autoregression techniques to measure spatial dependence and developing CNN + LSTM models for spatio-temporal feature learning and prediction.

Honors and Achievements

Autonomy Research Center For STEAHM Fellowship *2025 - 2026*
California State University, Northridge Dean's List, 5 semesters *2023 - 2025*
Moorpark College Dean's List *2022*

Personal and Class Projects

Canadian Climate Time Series Analysis and Clustering
- Conducted time series analysis on 60 years of climate data from 13 Canadian cities, applying SARIMA models to capture seasonal and trend components for accurate temperature forecasting
- Validated models through residual diagnostics and out-of-sample testing, achieving strong performance in forecasting and warming trend detection

Predicting Customer Dissatisfaction
- Built and evaluated classification models (KNN, LDA, QDA, Logistic Regression, Random Forest, XGBoost) on a 76,000-observation Santander dataset with 370 feature.
- Applied dimensionality reduction (PCA) and Random Forest feature selection to reduce high dimensionality, narrowing to 74 key predictors for improved interpretability.
- Tuned and validated models with stratified sampling, regularization, and grid search, achieving a high ROC-AUC score with XGBoost;
- Tools: Python, Jupyter Notebook, Anaconda, NumPy, Pandas, Matplotlib, scikit-learn, XGBoost.

Technologies

Languages: Python, R, C/C++, Java, JavaScript, TypeScript, SQL, HTML, CSS
Libraries/framework: React, Express.js, pydoc, pytest, NumPy, Pandas, Matplotlib, scikit-learn, XGBoost
Technologies: AWS (EC2, RDS), DigitalOcean, Node.js, Maven, Docker, Git, GitHub, JIRA, Jupyter Notebook, Anaconda
Systems Linux, Windows