STEVEN NING

Irvine, CA • steven20070@gmail.com • (669) 208-8389 • linkedin.com/in/stevenning123 • github.com/stevenning123

EDUCATION

University of California, Irvine *Expected Grad. June '28*

B.S. Data Science; B.S. Business Information Management ***GPA**: 4.00*

Relevant Coursework: Artificial Intelligence, Machine Learning, Computer Networks, Operating Systems, Data Structures & Algorithms, Linear Algebra, Probability & Statistics, Neural Networks and Deep Learning, Data Mining, Multivariable Calculus

EXPERIENCE

hireEZ ***Mountain View, CA***

Machine Learning Intern *Jun. 2024 - Aug. 2024*

- Improved an XGBoost classifier's F1 score by **5%** by building a Python and Pandas pipeline that deduplicated, normalized, and validated a **200,000-**row training dataset, reducing invalid, duplicate, and missing records by **18%** before model training
- Cut review time by **30%** for **20** recruiters by shipping **5** GPT-4o tool-calling agents that explain healthcare certifications in plain language, reaching **91%** accuracy on a **200-**example test set graded by rubric-based human review and prompt refinement
- Saved **~6** hours per week of manual data processing by automating ETL workflows for cleaned ML training datasets with Airflow, running **4** scheduled extraction, transformation, and loading jobs that delivered validated datasets to the ML team

Silicon Valley Cybersecurity Institute ***San Jose, CA***

Research Assistant *Jun. 2023 - Jul. 2024*

- Achieved **99.7%** accuracy, **99.53%** recall, and **99.87%** precision in network intrusion detection using **76%** fewer inputs (**10** of **41** features) by training a Random Forest on the top features ranked by entropy in a feature selection and model comparison study
- Compared **4** selection methods (entropy, mutual information, chi-squared, ANOVA) across **4** models (Random Forest, SVM, RNN, MLP) on **197,222** NSL KDD traces, with SVM accuracy ranging **94%** to **98.77%** while RNN and MLP varied under **1%** at **20** features
- Tuned each model on an **80:20** split with **10** and **20** feature subsets (Random Forest depth **10**, SVM RBF kernel with C of **100**, RNN with **50** LSTM units, MLP with **100** neurons), lifting RNN and MLP accuracy by up to **4%** moving from **10** to **20** features

Integem ***Cupertino, CA***

Lead Instructor Intern *Jun. 2026 - Aug. 2026*

- Taught **~300** students creative computing, Python, and generative AI at Camp Integem over **7** weeks, with every student building and presenting an original solo project each week, earning a **96%** camper satisfaction rating from **120** end-of-camp surveys
- Guided students to build a text-to-image generator in PyTorch using NVIDIA's Deep Learning Institute diffusion course (UNet, DDPM and DDIM sampling, classifier-free guidance, CLIP), so **90%** of campers generated and presented original AI artwork by week **4**
- Designed **7** original lesson decks covering Python, generative AI, 3D game design, Adobe Character Animator, and holographic AR, refining them weekly from student feedback across multiple cohorts and raising project completion from **70%** to **92%**

PROJECTS

PlayMatch, Sports Matchmaking App

- Built a sports matchmaking app with a React frontend and Express and PostgreSQL backend that pairs players by skill level, sport, and location through a layered REST API, matching **150** users into **60** games during testing with **85%** match acceptance
- Achieved **96%** backend test coverage across **30** Jest unit and Supertest integration tests by separating REST controllers, business logic, and data access into testable layers, with a written test plan covering scope, tools, and coverage targets

JaySchedule, AI-Powered Weekly Planner

- Built an LLM-powered planner on the Anthropic API that drafts **3** tone variants of email replies per Gmail thread using **6** prompt templates conditioned on thread context and desired tone for each reply, reaching a **78%** reply acceptance rate across **12** users
- Evaluated reply quality by comparing Claude and Gemini outputs on **50** real email threads using rubric-based human review graded on tone match, clarity, and accuracy, selecting Claude after it scored **15%** higher on tone match for the final reply drafts
- Cut weekly email and calendar triage time by **3** hours by building a serverless Netlify backend with Google OAuth that refreshes Calendar and Gmail data daily into a prioritized task list with week-grouped action items and completion percentages

PUBLICATIONS

S. Ning, K. Nguyen, S. Bagchi, and Y. Park, "The Study of Feature Engineering in Machine Learning and Deep Learning for Network Intrusion Detection Systems," 2024 Silicon Valley Cybersecurity Conference (SVCC), Seoul, South Korea, pp. 1–5.

SKILLS & INTERESTS

Languages & Frameworks: Python, JavaScript, TypeScript, SQL, C++, Java, R, HTML, PyTorch, TensorFlow, Keras, Scikit-learn, XGBoost, Hugging Face Transformers, NumPy, Pandas, React, Node.js, Express, FastAPI, Flask

Tools: Git, Docker, Linux, Jupyter, PostgreSQL, AWS SageMaker, GCP Vertex AI, OpenAI API, Anthropic API, Gemini API, Jest, Capacitor

Interests: Basketball, Marathon Running, Kaggle Competitions