

EDUCATION

- **University of California, Berkeley** Berkeley, CA
Master of Engineering, Data Science and Systems Aug. 2023 – May 2024
 - *Relevant Coursework:* Natural Language Processing, Machine Learning, Applications of Parallel Computing, Product Management, Technology Strategy, Team Management
- **Arizona State University (Barrett, the Honors College)** Tempe, AZ
Bachelor of Science, Computer Science; Minor, Korean Studies Aug. 2019 – May 2023
 - *Relevant Coursework:* Artificial Intelligence, Data Mining, Data Structures & Algorithms, Human-Computer Interaction (HCI), Database Management Systems

SKILLS

- **Programming Languages:** Python, C/C++/C#, JavaScript, SQL, Swift, Java, HTML, CSS
- **Libraries and Tools:** PyTorch, TensorFlow, scikit-learn, NumPy, Pandas, MySQL, PostgreSQL, React.js, Node.js, Git, CUDA, OpenMP, MPI, GitHub, GitLab, JIRA, JupyterLab, Hugging Face, BERT, GPT-2, Docker, JUnit 4, Maven, JSON, Swift, Linux, MacOS, iOS, Unity, Tableau, Kubernetes, Networkx, CMake
- **Skills:** Web Application Development, Full-Stack Development, Data Structures, Algorithms, Deep Learning, Machine Learning Models, Natural Language Processing, EDA, ETL, Data Processing, Data Visualization, Data Scraping, Model Evaluation, Feature Engineering, Fine-tuning, Object Oriented Programming, Mobile Development, Game Development, Agile, DevSecOps, CI/CD, SDLC, Bash Scripting, Parallel Computing, Clustering, Classification, Regression, Transformer Models, Large Language Models, Decision Trees, Multimodal LLMs, UI/UX, HCI, Unit Testing, E2E Testing, Database Design & Management, Software Architecture Design, Data Migration, Pipeline Development

PROJECTS & EXPERIENCE

- **Graduate Machine Learning Researcher** Berkeley, CA
UC Berkeley Capstone: Deep Learning for Patent Disambiguation September 2023 – Present
 - Collaboratively built a deep learning model coined as Patent2vec and utilized its output in a Decision Tree model, achieving 91% accuracy in classifying patents by inventor.
 - Utilized models from scikit-learn such as Random Forest and Logistic Regression to handle missing data, achieving an impressive 80% accuracy rate and enhancing overall data robustness.
 - Fine-tuned BERT, XLNet, and T5 encoders for title and abstract encoding, and compared performance.
 - Conducted feature engineering and data processing on patent data for model input.
 - Evaluated model performance using metrics like ROC curve.
- **Full-stack Software Engineer Intern** Tucson, AZ
Corsair Ranch Software Factory at Georgia Tech Research Institute May 2022 – Aug. 2023
 - Implemented full-stack features, including unit testing, database migrations, and models, for an original React.js web application enabling Air Force personnel to conveniently track pay records.
 - Rapidly acquired proficiency in web development and React.js during the internship, emerging as the team's go-to member for UI design.
 - Revamped company website security, user interface, and usability by developing a React.js-based platform. (<https://corsairranch.dso.mil/>).
 - Maintained Agile workflow with Jira and GitLab for weekly sprints.
- **Undergraduate Programming and Data Analysis Intern** Tempe, AZ
Center for Accelerated Operational Efficiency (CAOE) at ASU Aug. 2021 - July 2022
 - Under assistant professor Dr. Adolfo Escobedo, developed UI to gather training data for a computer vision tool to identify dangerous objects in airport baggage scans.
 - Implemented a mouse tracking feature on Amazon Turk survey, then visualized this user attention data in a heat map with NumPy to analyze and generate more effective training data.