

# Anastasia Alexandra Arsky

(650) 823-1087 • anarsky@gmail.com • [LinkedIn](#) • [Github](#) • Redwood City, CA

## EDUCATION

---

### UC San Diego, Extended Studies

Apr 2023 - Sep 2023

Machine Learning Engineering

### California Polytechnic State University, San Luis Obispo

Aug 2017 - Jul 2022

B.S. in Statistics, GPA: 3.4

## WORK EXPERIENCE

---

### Private Tutor

Sep 2022 - Present

Grade Potential Tutoring, Redwood City, CA

- Conducted personalized one-on-one sessions with students (ranging from elementary to high school levels) in various subjects, including Python and statistics.
- Provided guidance on classroom topics, assignments, and academic tasks.
- Developed supplemental curriculum for advanced learners.

### Inventory Manager

Mar 2020 - Jul 2021

Quail Creek Farms, Salinas, CA

- Utilized Metrc regulatory software and MS Excel to audit/manage inventory in compliance with state regulations.
- Web scraped the Metrc site to compile contact information of licensed companies, organized data into CSV files to create a marketing database.
- Created a comprehensive standard operating procedure handbook and conducted employee training.

## SIGNIFICANT PROJECTS

---

### Machine Learning Capstone Project for MLE Bootcamp: Spam Classification

Apr 2023 - Sep 2023

Tools: Python (Pandas, NumPy, Scikit-learn, Keras, TensorFlow, Xgboost), Jupyter Notebook, AWS

Merged the SpamAssassin and Enron Spam datasets; applied NLP techniques to classify email messages as spam or not.

- Explored traditional ML models (Random Forest, XGBoost) and deep learning architectures (BiLSTM, CNN).
- Conducted a comparative analysis of word embedding strategies: manual methods (TF-IDF, Continuous Bag of Words) vs transfer learning (pre-trained word embeddings, specifically GloVe, FastText).
- Scaled and deployed prototype to production environment using AWS SageMaker, Lambda, and S3.

GitHub: <https://github.com/anastasiaarsky/ML-Capstone-Spam-Classification>

### Data Science Project for Knowledge and Discovery From Data Course: Chat Bot

May 2022 - Jun 2022

Tools: Python (BeautifulSoup, Pandas, NumPy, PyMySQL), SQL

Collaborated as a team to implement a chat bot that answered questions about Cal Poly course information.

- Web scraped Cal Poly's course catalog and class offerings, populated a SQL database with the processed data.
- Utilized SQL queries to generate the chat bot responses.

GitHub: <https://github.com/jcavalca/bot-466>

### Data Science Summer Project: Knowledge Graph Service

Jul 2021 - Aug 2021

Tools: Python (Pandas, Asyncio), SQL, Presto, Hive, Mercurial, Phabricator

Independently developed algorithms using NLP techniques to improve a knowledge graph search engine.

- Implemented NLP algorithms for synonym generation (for the concept graph) and common compound word identification amongst the named entities (to refine search).
- Designed algorithm to detect anomalies in knowledge graph builds.
- Conducted extensive testing to identify commonalities between failing cases in the search engine.

## SKILLS

---

Python, data science, machine learning, statistical modeling, data analysis, deep learning, neural networks, statistics, data preprocessing, feature engineering, natural language processing, data visualization, SQL, Java, Git, AWS