SAFIYA ALAVI

THIRD-YEAR STATISTICS AND DATA SCIENCE

408 – 786 – 8918 ☎
safiyazalavi@ucsb.edu ✉
linkedin.com/in/safiyaalavi 🔗
github.com/safiyaalavi02 🔒

TECHNICAL EXPERIENCE

DATA SCIENCE CAPSTONE
Amgen | Sept 2022 – Present
- Created **knowledge graph** using **NLP** and **text mining** algorithms to increase efficiency of searching for entities and their relations, discovering new information and enabling complex decision making
- Quantified performance of existing NLP algorithms used for knowledge graph generation (**Named Entity Recognition and Linking** and **Relation Extraction**) by testing against generated benchmark dataset containing labels for entities which relate to the biomedical field

INDEPENDENT RESEARCHER
Moeller Lab | Sept 2022 – Present
- Collaborated with professors at **UCSB** and **MIT** to analyze effects of global warming on carbon and nutrient cycling in the ocean through use global-scale ecosystem models
- Produced **visual outputs** of simulated data by local-scale and **global-scale models** using **MATLAB** to visualize how the thermal sensitivity of plankton metabolism effects nutrient cycling world-wide

META DATA CHALLENGE FINALIST
Meta | May – Aug 2022
- Selected as a finalist (**one of eighty-six** undergrad students in North America) for a competitive **five-month** data science apprenticeship at Meta
  - Involved small and large group projects, one-on-one mentorship opportunities, peer networking and technical workshops
- Analyzed Netflix Kaggle dataset using **R-Studio** and **SQL** to produce actionable insights for a **business strategy** and presented a **data-driven product pitch** to an audience of Data Scientists and Engineers

UC DAVIS COSMOS ATTENDEE
COSMOS | Jun 2019 – Jul 2019
- Earned invitation to UC Davis COSMOS, a prestigious five-week program (**30% acceptance**) revolving around key biology and coding concepts
- Collaborated with peers to study, outline and create a presentation on the evolution of antibiotic resistance which was showcased to an audience of over **five-hundred people**

PROFESSIONAL & LEADERSHIP EXPERIENCE

UCSB DHADKAN PRESIDENT
UCSB Dhadkan | Sep 2020 – Present
- Directed non-profit organization of **forty+ students** and seven sub committees to generate over **$10,000 in profit** by facilitating small and large fundraisers throughout the year
- Managed the marketing, financial, logistical and social facets of Nachle Deewane, an annual dance competition which attracts over a **thousand attendees** from around the US
  - Involved mass communication, overseeing tasks for 9 executives, planning advertising campaigns

MATHEMATICS TUTOR
Mathnasium | Jan 2020 – Nov 2021
- Enhanced over **fifty students'** capabilities and self-esteem in elementary to **university-school level mathematics** with hands-on, personalized tutoring styles and worksheets customized to their learning
- Communicated with fellow instructors to ensure students received a wide variety of tutoring styles

EDUCATION

UC Santa Barbara
Candidate for Bachelors in Statistics and Data Science
Honors Student
GPA: **3.89/4.0**
Relevant Courses:
- Statistical Machine Learning (92%)
- Design and Analysis of Experiments
- Regression Analysis (97%)
- Applied Stochastic Processes (92%)

TECHNICAL SKILLS

LANGUAGES:
- Python (pandas, numpy, matplotlib, pyspark)
- R-Studio (tidyverse, dplyr, ggplot)
- MATLAB
- SQL
- SAS
- Microsoft Excel

PROJECTS

Wine Quality Classification (Supervised ML Project in R-Studio)
Identified and predicted wine quality given chemical predictors using ML models such as LDA, Single Decision Trees, Random Forest, Boosted Trees

"ZuckFlix" Content Recommendation
Provided actionable information and created intuitive visualizations about current trends in the entertainment industry using SQL and R

AWARDS

- L&S Honors Program
- AP Scholar with Distinction

HOBBIES & INTERESTS

- Reading
- Hiking
- Travelling
- Photography
- Chess