Tanya Beri

1521 Greene Dr, San Jose, CA 95129 | (408) 930-5593 | tberi2003@gmail.com | www.linkedin.com/in/tanya-beri

EDUCATION

UNIVERSITY OF CALIFORNIA, LOS ANGELES Los Angeles, CA

Bachelor of Science, Statistics and Data Science Major, Mathematics Minor Expected Graduation June 2024

- Relevant Courses: Programming in C++, Data Structures, Algorithms, Statistical Programming in R, Linear Algebra, Differential Equations, Discrete Mathematics, Multivariable Calculus, Statistical and Data-Driven Modeling, Regression

SKILLS

- Technical Skills: Proficient in Python, R, C++, SQL; Familiar with Java, AI and ML; Research and Data Analysis
- Soft Skills: Communication, presentation, teamwork and collaboration

EXPERIENCE / PROJECTS

Denodo Technologies Inc, Palo Alto, CA Jun 2023 - Sept 2023

Data Science Engineering Intern

- Optimized Python code to gather statistics from customer views and queries in Denodo Platform to assess their performance
- Designed and created an executive dashboard using Apache Superset and SQL to provide clients with insight on performance for data being stored in Denodo Platform and highlight bottlenecks to help with troubleshooting and optimizing performance
- Created Python script to seamlessly integrate new data into dashboard and reflect growth in performance over time
- Used VQL (Virtual Query Language) to gather information on cache jobs and display metrics about cache performance

UCLA, CA Sept 2022 - Dec 2022

Project Member

- Completed a project to build a CO_2 emissions forecast model for Data-Driven Mathematical Modeling course at UCLA
- Implemented Random Forest Classifier to test for feature importance and KNN algorithm for multi-class classification of CO_2 emissions by country

UC Santa Cruz, CA Jun 2022 - Aug 2022

Data Science Research Intern

- Used TCGA-PAAD (Pancreatic Adenocarcinoma) dataset to analyze accuracy of RNA quantification tool MESA
- Cleaned data and created 2 PCA plots to analyze differences in the normal and tumor data samples
- Determined the splicing signature of the tumor type and detected and further researched a set of outlier samples that were clustering separately in the PCA plots

UC Davis, CA Jun 2019 - Sept 2019

Software Engineering Research Intern

- Determined the protein domains found in each gene from the A. crassicarpa genome (Acra_v1) using BLAST
- Created a Python script to parse through the BLAST output and detect patterns, outputting a table with the gene ID and protein domain information

San Jose State University, CA Jun 2019 - Sept 2019

Data Science Independent Research Intern

- Used Python to create a K-Means clustering model of lung cancer data to show which factors were most common among lung cancer patients
- Created and presented weekly slides along with three lab and project reports detailing the methods and algorithms used

EXTRACURRICULARS

Data Resolutions at UCLA, Project Member Sept 2022 - Present

- Worked with client OKTY to create a model to recommend specific advertisements based on user's spending patterns and presented model and methods to the client
- Used Named Entity Recognition to anonymize transaction data to protect user privacy

ROOT|D Dance Team at UCLA, Member Sept 2021 - Present

- Learning various styles of dance and performed at UCLA Garba, USC Mehfil, and St. Baldrick's PULSE fundraiser
- Planned and hosted "Mock Mehndi" event at UCLA with 200+ attendees and 10+ performances