Sidharth Babu (CV)

LinkedIn : https://www.linkedin.com/in/sidharth-babu
Github : https://www.github.com/SidharthBabu121
Email : sidharthbabukish121@gmail.com
Phone : (408)-649-0393

EDUCATION

University of California, Santa Cruz **September 2022 - Current**
Master of Science, Natural Language Processing

University of California, Merced **July 2018 - May 2022**
Bachelor of Science, Computer Science and Engineering(CSE)
GPA: 3.7

HONOURS AWARDS

- Chancellor's Honour List **2019-2021**
- Innovate to Grow Fall : Winter 2019 : Top Finisher **2019**
- Dean's Honour's List **2018-2021**

TECHNICAL SKILLS

Programming Languages: Java, C/C++, Python, SQL, JavaScript, HTML,CSS,R, MathLab
Frameworks/Libraries: Pytorch, TensorFlow, Keras, BERT, Pandas, NumPy, Matplotlib, AstroPy, SciPy
Developer Tools: Git, Docker, Amazon Web Services (AWS), VS Code, Visual Studio, PyCharm, Eclipse

TECHNICAL EXPERIENCE

Summer TITANS Research Intern **May 2022 - August 2022**
Sandia National Laboratories

- Applied statistical modeling using Pandas, Matplotlib, Seaborn and SciPy to perform ad-hoc analysis on Sandia's commercial-off-the-shelf (COTS)program before presenting correlations and relationships among measured variables to over 20 clients.
- Utilized exploratory data analysis techniques to compare customer needs for over 8 different 3rd party Sandia affiliated sites to existing datasets and collected insights by generating appropriate data visualizations for new use case opportunities.
- Improved data analysis efficiency for ad-hoc analyses by 90% through supervised learning into an existing Python software framework to enable scalability and reusability.

Data Science Challenge Summer Intern ***May 2021 - June 2021***
Lawrence Livermore National Laboratory

- Developed and optimized supervised and unsupervised machine learning models with over 95.4% precision and 96.3% recall rate using Numpy and AstroPy on recent time-domain optical astronomy data to detect, distinguish, and characterize Near Earth Objects (NEOs).
- Developed highly refined Python scripts to categorize, capture, store, and share provided astronomical data in real-time, increasing data throughput and data accuracy by 15% with cohort.
- Delivered a 25 minute presentation to 50 LLNL Team members detailing findings and potential applications of the developed predictive models and their corresponding solutions.

Joint Genome Institute (JGI) Summer Intern ***June 2020 - August 2020***
Berkeley Labs

- Evaluated and optimized a hybrid clustering strategy with the bio-python library on the Lawrencium Cluster before processing and clustering reads using a hybrid strategy to produce an assembly that contained 42 contigs and 30,000 reads.
- Leveraged and helped test a revolutionary scalable metagenomic clustering tool called SpaRC on the complementary characteristics of both short and long-read sequencing platforms to generate taxonomic profiles with 94% accuracy compared to existing databases.
- Developed and tested pipeline to analyze, process and visualize data from Illumina short and long reads to predict bacterial abundance, increasing accuracy by 33% with minimally reduced runtime.

PROJECTS

Clickbait Spoiling via Spoiler Generation ***September 2022 - January 2023***

- Researched and implemented a natural language processing model to automatically generate spoiler sentences to mitigate clickbait through a transformer based extractive QA model that outperformed the baseline by 16.3% in accuracy.

- Experimented with 3 models, 2 using vanilla BERT for sequence classification and the third using a combination of BERR, and BiLSTM for sequence classification before applying various hyper-parameters tuning techniques to improve context-based word understanding accuracy by 5% using a modified BLEU F1-Score

Project Protect: Healthy Host ***August 2019 - December 2019***

- Implemented the Hmong, English and Spanish Section using React Native in a multi-platform app while updating a maintenance manual for non-coders with the Content Team to facilitate new markets and better serve the existing 500+ client base to improve access to healthcare for the multi-ethnic communities in Merced County.
- Competed in the community wide Fall 2019 Innovate to Grow Competition at UC Merced and was Issued a certification and cash prize of over $2,500 in recognition as the top finisher.

LEADERSHIP ACTIVITES

Co-Founder and Treasurer ***August 2018 - August 2019***

Electrical Engineer's Society of UC Merced(EES)

Interfaced with ASUCM to secure over a $1900 for projects to spark and guide people's interest in science and technology while engaging in professional networking.

ASSOCIATIONS

- Association of Computing Machinery (ACM) 2018-2022
- UC Merced Robotics Society 2019-2022
- Society of Asian Engineers and Scientists(SASE) 2019-2022
- The Foster Family Center for Engineering Service Learning(ESL) 2019-2022
- Institute of Electrical and Electronics Engineers (IEEE) 2018-2022