SIDHANT PUNTAMBEKAR
Computer Scientist | Computational Biologist

sidhantnp@yahoo.com   https://www.linkedin.com/in/sidhant-puntambekar/
720-250-8098   https://github.com/SidhantPuntambekar

Education

Bachelor of Arts in Computer Science, Computational Biology Minor, Molecular Biology Minor
University of Colorado Boulder, Boulder, CO, August 2019 - May 2023
- University of Colorado Esteemed Scholar. **Overall GPA: 3.964, Major GPA: 3.979**

Work Experience

Computational Biology/Data Science Intern, Multi-Cancer Early Detection (MCED) Team
Exact Sciences Corporation, Boston, MA, *May 2022 - August 2022*
- Conducted statistical analysis in Python and R for accuracy improvements of genomic mutational variant biomarker signal in CancerSEEK, a developmental liquid biopsy based diagnostic test for early detection of stage I and II cancers.
- Analyzed effects of unique molecular identifier (UMI) contamination on 40 genomic primer targets in the SafeR-SeqS duplex sequencing assay to improve sensitivity and specificity of a rare variant machine learning based logistic regression classifier.
- Leveraged Apache Spark and Microsoft Azure Databricks to quantify UMI contamination from over 430,000+ sequencing reads.

Biomedical Informatics Research Intern, Dr. Shamil Sunyaev Computational Genomics Lab
Harvard Medical School, Boston, MA, *May 2021 - May 2022*
- Analyzed the Broad Institute's Genome Aggregation Database (gnomAD) dataset to better quantify and understand the mutational constraint on missense and loss of function genomic variants using a fine-scale mutational map.
- Modeled baseline mutational spectrum using machine learning techniques (logistic regressions) on a training set of over 125,000 synonymous genome and exome sequences leveraging scikit-learn, pandas, and dask.
- Built a logistic regression classifier to predict genes that are more likely to succumb to debilitating loss of function mutations.

Tech Team Director, HackCU
University of Colorado Boulder, Boulder, CO, *May 2020 - Present*
- Developed the new HackCU hub website (hackcu.org) utilizing full stack development tools and continuous integration (CI/CD) as part of the technology/software development team.
- Migrating HackCU website over to Next.js framework to integrate backend, database, and content management system.
- Hosted workshops during hackathons on bioinformatics and recruited new contributors to HackCU Development team.

Bioinformatics Research Intern, RNA Bioscience Initiative
University of Colorado Anschutz Medical School, Denver, CO, *May 2020 - May 2021*
- Authored (primary authorship) research paper published in Public Library of Sciences Biology journal (May, 2021) regarding the problem of single cell RNA-sequencing metadata submission in gene expression databases (NCBI GEO).
- Developed RShiny web application for open source 'ClustifyR' R library in order to facilitate analysis of scRNA-seq data.
- Created a meta-analysis of 15 scRNA-seq data sets from NCBI Gene Expression Omnibus and Tabula Muris (mouse scRNA-seq atlas) utilizing principal component analysis, graph based clustering, and uniform manifold approximation and projection.

Projects

Human-Symptom Disease Network (HSDN) Community Detection
- Evaluated various community detection algorithms such as the degree corrected stochastic block model (DC-SBM) and Louvain clustering to elucidate disease-symptom phenotypic relationships in the HSDN network by Zhou et al., 2014.

Sorry! Board Game
- A Java rendition of the board game "Sorry!". Built using JavaFX, JUnit testing, and object oriented design patterns including factory, command, observer, and model-view-controller. Employed relational MySQL database to log completed game results.

Skills & Technologies

Languages: Python, Java, C/C++, R, HTML5, CSS, JavaScript, TypeScript, SQL, MATLAB, Scala
Testing Tools: Mocha, Chai, Jest, Cypress, JUnit
Continuous Integration Tools: TravisCI, CircleCI, Azure DevOps, Jenkins
Production Tools: Google Cloud Platform, AWS, Azure, Databricks, Apache Spark, Git, Heroku, Docker/Kubernetes, NPM, Yarn, Surge
JavaScript Technologies: ReactJS, NodeJS, ExpressJS, PassportJS, Bootstrap
Databases: MongoDB, MySQL, PostgreSQL
Operating Systems: Mac OS X, Linux (Ubuntu, CentOS, Debian), Windows