Venkatesh Kamaraj

venkatk89.github.io
github.com/venkatk89
vengatesh.vengatesh89@gmail.com

Research Experience

Project Scientist Jun 2022 – Until now
Junior Research Fellow Jan 2021 – May 2022
Centre for Integrative Biology and Systems Medicine, IIT Madras Chennai, India
Advisors: Dr. Himanshu Sinha, Dr. Karthik Raman, Dr. Manikandan Narayanan

Key Projects

- **Analysis of human genomes with genome graphs**
  - Designed scalable computational workflows with multiple bioinformatics tools for WGS studies
  - Developed novel methods for genome graph annotation to identify regions of significance
  - Performed detailed examinations to study the structural complexities in the genome graphs
- **GenomeIndia: Unveiling unique variants in Indian (sub) populations**
  - Analyzed the WGS of individuals from India's diverse populace using genome graphs
  - Used the dynamic nature of the genome graph to capture prevalent and rare variants in Indians
  - Studied the consequences of genetic variations within and between Indian sub-populations
- **Approachability of genomics and visualization of variants**
  - Ideated and developed SCI-VCF, a cross-platform *application* for genomic data analysis
  - Designed an intuitive GUI for the tool to aid users irrespective of their programming expertise
  - The software helps users summarise, compare, and visualize genetic variants from VCF files
- **Polygenic Risk Scores for common complex diseases**
  - Computed individual-level genetic risk scores for type-2 diabetes in an Indian cohort
  - Assessed the effect scores from homogeneous and trans-ethnic GWA Studies as base datasets
  - Employed machine learning techniques to enhance the predictive power of the calculated scores

Education

Bachelor of Technology in Engineering Physics Aug 2015 – May 2019
Indian Institute of Technology, Madras | *CGPA: 7.18/10* Chennai, India

Publications

- Venkatesh Kamaraj, Himanshu Sinha, SCI-VCF: a cross-platform GUI solution to summarize, compare, inspect and visualize the variant call format | NAR Genomics and Bioinformatics, 2024 | doi
- Venkatesh Kamaraj, Ayam Gupta, Manikandan Narayanan, Karthik Raman, Himanshu Sinha, Unveiling Genomic Complexity: A framework for genome graph structural analysis and optimised variant calling workflows | bioRxiv, 2024 | doi

Presentations

A Deep Dive into Genome Graphs: Structural Implications and Variant Calling Workflows May 2024
Poster presentation at the Inaugural WSAI Annual Research Showcase

GenomeIndia: Cataloguing the genetic variations in Indians Aug 2023
Poster presentation at the RBCDSAI AI/ML conclave on healthcare

Sequence graph representations of yeast and human genomes April 2021
Talk at the Data Science workshop organized by the GenomeIndia consortium

Conferences and Training

IITM-BioModels Workshop in collaboration with EMBL-EBI Aug 2024
Teaching Assistant for the construction of reproducible ML models in Systems Biology

Data Science-driven solutions to improve maternal and child health Feb 2023
Reviewed the development of pregnancy dating models and challenges in clinical translation

Clinical Genomics to Systems Medicine: Transforming Healthcare Feb 2022
Focused on the novel multi-omics analysis methods for understanding complex human diseases

Data Scientist Career track with R – DataCamp Sep 2018
Gained in-depth expertise in the multifarious aspects of data science to interpret complex data

Work Experience

Data Science Consultant Sep 2019 – Feb 2020
Indus Insights and Analytical Services Gurgaon, India

- Developed recommendation engine for the biggest airlines in the USA, using cutting-edge ML
- Leveraged cloud computing to train and evaluate deep learning models on over 100 GB of data
- Audited the marketing models of a US-based small business lender by performing detail-oriented examinations on the SAS-based models to ensure robustness during deployment

Data Science Intern May 2018 – Jul 2018
Blitzkrieg Retail Private Limited Chennai, India

- Developed a recommendation engine for the online pharmacy store by mining association rules
- Created a proprietary image processing application for the company from scratch, using OpenCV
- Built an end-to-end ETL pipeline on Firebase and MSSQL to develop the company's dashboard

Other Works

DNA Sonification Tool Oct 2020

- Incorporated music theory and developed an auditory display *tool* to sonify genomic sequences
- Enhanced the musicality of the tool's output while maintaining its overall analytical capabilities

DengAI – Disease Spread Prediction Jul 2020

- Designed ensemble methods to forecast the spread of dengue using TSA and statistical modeling
- Trained the models on Collab GPUs with data collected over a decade in South American cities

Climate Data Analysis – Department of Chemical Engineering, IITM Aug 2019

- Analyzed the data collected across India and validated key insights through hypothesis testing
- Created temporal plots and geospatial heat maps for radiation and air quality-related parameters

Fraudulent Transaction Prediction – Department of Computer Science, IITM Apr 2018

- Conceptualized and designed an online fraud prediction algorithm using random forest classifiers
- Won the first position in the Exebit Data Science Challenge, 2018

For more details, visit *portfolio website*

Skills

Bioinformatics: Genome analysis, Single-cell transcriptomics, Statistical methodology, High performance computing, Data visualization, Machine learning, Deep learning, Software development

Programming: R, Python, SQL, C, HTML, Shell scripting, Snakemake, Docker, Git