Naveen **Kannan**

RESEARCH ASSOCIATE

Cleveland, Ohio

☎ +1 973 519 5397 | ✉ naveen.kannan12@gmail.com | 🏠 naveenkannan.netlify.app | ⓞ naveenk2022 | in naveen-kannan-177120118

Data Scientist experienced with data analysis, visualization, building scientific pipelines and Data Engineering. I have a strong passion for learning, documentation and automation, and I compiled this CV with code. United States Permanent Resident.

Education

MSc in Biomedical and Health Informatics
Cleveland, Ohio

CASE WESTERN RESERVE UNIVERSITY
JANUARY 2022 – AUGUST 2023

- GPA-3.9

Bachelor of Medicine and Bachelor of Surgery
Chennai, India

MADRAS MEDICAL COLLEGE
SEPTEMBER 2016 - MARCH 2020

- GPA-4.0

Experience

Research Associate
Cleveland, Ohio

POPULATION HEALTH AND QUANTITATIVE HEALTH SCIENCES, SCHOOL OF MEDICINE, CASE WESTERN RESERVE UNIVERSITY
MAY 2023 - Present

- **Docker/Singularity Container Deployment and Documentation**
    - Developed Docker/Singularity containers for reproducible scientific pipelines, facilitating efficient workflows for colleagues.
    - Authored comprehensive documentation to ensure seamless adoption and easy troubleshooting.
    - Incorporated Mamba into the scientific workflow, reducing container build times by 300 percent.
- **Cluster Configuration and Big Data Analytics**
    - Set up, configured and maintained both VM and Raspberry Pi bare-metal servers from scratch.
    - Created in-depth documentation for the cluster set-up and configuration.
    - Implemented big data analytics tools (Hadoop, Spark) along with bioinformatics tools like Hail and Glow using Ansible scripts.
- **Genomic data Analysis**
    - Collaborated with the Alzheimer's Disease Sequence Project by annotating and scoring the biological effects of 362 million structural genetic variants associated with Alzheimer's Disease.
    - Performed statistical analysis on the data and presented findings in the form of a poster at the ASHG (American Society of Human Genetics) 2023 Conference in Washington DC.

Graduate Research Assistant
Cleveland, Ohio

POPULATION HEALTH AND QUANTITATIVE HEALTH SCIENCES, SCHOOL OF MEDICINE, CASE WESTERN RESERVE UNIVERSITY
JANUARY 2023 - MAY 2023

- Engineered Docker containers for streamlined deployment of specialized R packages and Python modules, optimizing analyses for single-cell RNA datasets.

Graduate Research Assistant
Cleveland, Ohio

GIS HEALTH AND HAZARDS LAB, SCHOOL OF MEDICINE, CASE WESTERN RESERVE UNIVERSITY
MAY 2022 - DECEMBER 2022

- **Mapping human movement patterns after a natural disaster.**
    - Developed Python scripts automating GPS coordinate extraction from GPX files of geospatial video feeds in Democratic Republic of Congo the volcanic eruption of Mt Nyiragongo.
    - Processed and transformed data into machine-readable format, enabling streamlined map generation.
    - Applied automated scripts to parse video feeds into frames, implementing Machine Learning Algorithms for tent detection.
    - Created heat maps of tent distribution according to GPS coordinates, allowing for the mapping of the human movement pattern in the aftermath of the eruption.
    - Wrote up the findings as a part of a manuscript for a peer-reviewed publication.

Junior Resident Doctor
Chennai, India

DEPARTMENT OF PSYCHIATRY, SAVEETHA MEDICAL COLLEGE
JULY 2020 - SEPTEMBER 2020

- **Diagnosing and managing patients with psychiatric illnesses**
    - Learned to evaluate and diagnose psychiatric illnesses in stressful environments.
    - Assigned and achieved a goal of management of five patients per day.

Junior Resident Doctor
Chennai, India

MADRAS MEDICAL COLLEGE
MARCH 2019 - MARCH 2020

- **Rotated through the following departments:** Internal Medicine, General Surgery, Obstetrics and Gynecology, Community Medicine, Psychiatry, Emergency Trauma Ward, Labor Ward.

Technical Skills

Operating Systems

Unix, Kubuntu Debian Linux, Fedora Red Hat Linux, Raspbian Debian Linux, Windows 11

Programming Languages

Python, R, Bash, SQL

Software

Hadoop/Spark/Hive, Jupyter, Docker, Singularity, SLURM, Ansible

Geospatial Analytics/Visualization

geopandas, sf (R), QGIS

Markup Languages

LaTeX, Markdown, RMarkdown

Version Control

Git

Python-based skills

PySpark for Hadoop/Spark, Tensorflow, Pytorch, pandas, scikit-learn, numpy, scipy

Bioinformatics tools

Samtools, BCFtools, HTSLib, Tabix, Plink