Nishant Gerald

880 W Peachtree St NW, Apt 1505
Atlanta, GA 30309
nishant.gerald@gmail.com
424-333-4638

EDUCATION

Georgia Institute of Technology, Atlanta, GA — **Graduation: Dec 2019**
M.S., Bioinformatics — **GPA: 3.92**

Manipal Institute of Technology, Manipal, India — **Graduation: May 2018**
B.Engg. in Biotechnology Engineering, Genetic Engineering minor — **GPA: 3.82**

EXPERIENCE

Data Scientist — **Jun 2021 – Present**
Global Payments Inc., *Atlanta, GA*

- **Sales Forecasting Engine:**
  - Spearheaded the first ever machine learning efforts within the Analytics and Customer Engagement business unit, that will benefit over 500,000 merchants across the USA and Europe.
  - Created and maintained the sales forecasting machine learning algorithm using an ARIMA framework to use historical sales data to make merchant sales predictions with an accuracy of 80%.
  - Improved on the older version of the sales forecasting engine using a Prophet model by including regional holidays and weather as features - bumping up accuracy to 90%.
  - Developed the data pipelines, along with the machine learning pipelines that are deployed on Google Cloud Platform.
- **Business Intelligence:**
  - Worked alongside business teams to set up complex SQL queries to aggregate and transform transaction level data, that helps our team better understand how the business is running.
  - Set-up monitoring dashboards for our business to track meta-level information easily
- **Team growth:**
  - Lead our data science team and manage one other data scientist.
  - Involved in the hiring of multiple engineers, product managers and project managers.

Bioinformatics Engineer — **Jan 2020 – Jun 2021**
General Dynamics Information Technology, *Atlanta, GA*
(Centers for Disease Control and Prevention - Scientific Computing and Bioinformatics Support Contract)

- **Linux System Administration:**
  - Troubleshot infrastructure-level issues for applications and services on 100+ CentOS systems.
  - Installed 150+ Unix-based applications from source which were configured with an LMOD module system and set-up scripts via Jenkins to automate the upgrades of several programs including Python, R and sratoolkit.
  - Developed LDAP automation pipelines including a GUI application for simpler and (18x) faster user onboarding into the CDC SCBS environment.
  - Deployed 50+ CRON jobs for system monitoring and alerts.
  - Supported the management of Samba ACLs for 100+ groups at CDC.
  - Worked with multiple vendors to help integrate commercial products into CDC's Linux environment.
- **High Performance Computing (HPC) Cluster Administration:**
  - Supported the management of CDC's high performance computing cluster comprising 80+ nodes and 2000+ cores which was managed using a Univa Grid Engine job-scheduler.
  - Deployed tools for simpler cluster-status and job monitoring
  - Provided support to users at CDC by identifying bottlenecks in cluster-jobs and customizing their submission scripts to optimize job-performance.
- **Website Development:**
  - Lead developer of the Scientific Computing Team's new intranet website to host training content and user-guides – hosted on an Apache web server with a MariaDB database.

- Set up Google Analytics reporting to track client traffic and metrics such as user sessions, tile-clicks and page visits to make data-driven decisions for user-training programs.

- **Technical Instructor:**
  - Developed and presented training sessions for 2200+ users at CDC on topics involving the use of Python for Data Cleaning, Analysis and Visualization, and pipeline development through over 20 in-person and remote webinar sessions.
- **Helpdesk Management:**
  - Managed the helpdesk using the Request Tracker (RT) ticketing system.
  - Directly handled and resolved over 1400 client tickets spanning issues including data management and access, code review, and containerizing applications.

Graduate Research Assistant **Aug 2018 – Dec 2019**
Lachance Lab, Georgia Institute of Technology, *Atlanta, GA*
Supervisor: Joseph Lachance, Ph.D.

- **Pipeline Development:** Developed STRUCTUREpainter – a parallelizable bioinformatics tool to paint genome-wide estimates of human population-structure for each of the 23 chromosomes.
- **Performance Optimization:** Optimized the data-processing parameters to strike a balance between computational speed and accuracy to speed up the process by 100%.

Genomics Data Scientist - Intern **May 2019 – Aug 2019**
Ancestry.com LLC, *Lehi, UT*

- **Pipeline Development:** Developed a variant detection pipeline that analyzes Next Generation Sequencing (NGS) data from customer DNA samples and identifies and highlights structural variations.
- **Quality Control (QC) Tool:** Developed a Quality Control application using Python and Flask to visualize and summarize Next Generation Sequencing data to easily identify faulty samples and reduce rerun costs.

Undergraduate Research Assistant **Jan 2018 – May 2018**
Saini Lab, Indian Institute of Technology, *Mumbai, India*
Supervisor: Supreet Saini, Ph.D.

- **Computational Model:** Developed a computational model to simulate a biological mutation network of over 100 thousand nodes using Python and ran an optimization algorithm on this network to visualize the fitness landscape of the bacteria.
- **Publication:** First author on the resulting publication in Springer-Nature's Systems Biology Journal

PROGRAMMING LANGUAGES AND TECHNOLOGIES

- **Scripting Languages:** Python (highly proficient), Bash (highly proficient), PHP and JavaScript.
- **Operating Systems:** Windows, macOS and Linux (Ubuntu, RHEL, CentOS)
- **Cloud Computing Technologies:** AWS – EC2, S3, Lightsail, IAM and RDS
- **Other Technologies:** UGE Cluster Computing, Snakemake, Docker, Singularity, LDAP, Tableau, HTML, D3.js, PHP, SQL, MariaDB, Samba ACLs, Jenkins, Web Servers (Apache and Nginx), Google Analytics and Google DataStudio

PUBLICATIONS

- Mathematical modeling of movement on fitness landscapes (2019): **Gerald, N.**, Dutta, D., Brajesh, R. G., Saini, S. (2019). Mathematical modeling of movement on fitness landscapes. BMC Systems Biology,110. https://doi.org/10.1186/s12918-019-0704-0

HONORS AND AWARDS

- **Computational Biology Faculty Research Awards for 2018 and 2019:** Awarded for excellence in research and academics across both Fall and Spring Semesters of 2018 and 2019.

INTERESTS

- At-home technology projects - currently working on a motion-sensing tool

- Beer-tasting at local breweries
- Soccer (represented at both school and collegiate levels)