Ketan Jog

kj2473@columbia.edu | github: ketanjog | 925-316-5709
9570 Columbia Student Mail, New York, NY 10027

Education

Columbia University — New York, NY
Bachelor of Arts in Math-Stats, B.S Computer Science; GPA: 3.6/4.0 — *Expected May 2023*

Currently: Junior (*Designated Science Research Fellow)*
Programming Languages: Python, C, C++, R, Java, HTML, CSS

Experience

Dr George Dragomir — Columbia University
July 2021 - Ongoing

Worked on Mathematically modelling miner-blockchain interactions in Bitcoin-like blockchains. Further designed a simulator that tested the robustness of blockchain based systems against deviant mining attacks

Dr Nakul Verma — Columbia University
Independent Research — *Jan 2021 - May 2021*

Working on a theoretical unsupervised learning problem in sample-complexity analysis. Aim to explore the stability of the embedded space generated by dimensionality reduction algorithms while sampling from an underlying distribution with specific properties.

BinIt Inc. — Columbia University
Director and Co-founder — *Oct 2018 - present*

Co-founded BinIt, an analytics and resource management software company for waste management. Worked on Object localisation and detection pipelines using computer vision techniques. Wrote lightweight algorithms to be run on low-spec systems. Built prediction algorithms using proprietary data. Experience in micro-service architectures and using google cloud services.

Mski Lab — Memorial Sloan Kettering Cancer Centre
Research Intern — *June 2020 - July 2020*

Contributed to a Precision Oncology project called *dryclean* - a signal processing algorithm that learns biological and technical noise in read-depth data and reduces the noise in signal for tumor samples, thus enhancing *CNA/CNV* detection, which is imperative for personalised cancer treatments. I worked on code reproducibility. Program was cut short due to Covid crisis.

Shiu Lab — Michigan State University
Research Assistant — *May 2019 - July 2019*

Completed iACRES, a Prestigious Computational Research Training Program. Placed into Shiu Lab to develop and use Deep Learning Methods to analyze Genomic Data and predict gene expression under environmental stress. Presented work at the MidSURE Symposium. Poster here

LionBase — Columbia University
Executive/Data Scientist — *May 2019 - present*

Worked on client acquisition and product development for LionBase - a student-driven LLC that creates data solutions for industry problems. Transitioned to Executive Team role in Dec 2019. Won the Columbia Venture Competition 2020 and received a $18,000 grant to develop the company. Filing for dissolution in 2021 due to Covid stall.

Summer Science Program in Biochemistry — Purdue University
Research Student — *June 2017 - Aug 2017*

Worked on a mentored research project with a team to design a fungal enzyme inhibitor with Wet Lab techniques as well as Computational Modelling and Simulation software. Extensively used Bioinformatics techniques and successfully published a paper.

Data Product Initiative — Columbia University
Data Scientist and Board Member — *May 2019 - January 2020*

Served on the Board of the Data Product Initiative - a Columbia user focused club that works on real-world data product ideas. Also worked on developing a support tool for novice investors to build portfolio management skills

Achievements & Projects

- As part of a class on unsupervised machine learning, conducted an experimental study of UMAP and tSNE to explore quality and stability of embeddings under various sampling contraints. Code here.
- As co-founder of BinIt, won the Columbia 2018 Design Challenge, and conducted a fully funded 2-week customer discovery trip to Mumbai, India. Won the Columbia Venture Competition 2019 and received a $25,000 grant and $100,000 of computational credit. Part of Almaworks, a non-profit Business Accelerator.
- Developed a sentiment analysis portal to monitor changes in nuanced sentiment in response to the Covid19 crisis. Portal here. Code here. Got mentioned at the Columbia-Covid Symposium
- Contributed to a project that developed linear models to predict high level energies of molecules using low level theory at chemical accuracy. Work involved feature engineering for linear regression models
- Placed into the top 7 finalists at the **Columbia CDSS Data Hackathon 2019** for project that predicted E-commerce adoption potential based on grocery trends. Won the Wolfram Award.
- Got into highly selective Berlin Core Curriculum Program for Summer 2019. Couldn't attend due to financial reasons.
- Won Second Place at the **Columbia CSI Space Hackathon 2019**. Presented a design concept solution for space debris tracking and collection.
- Completed the learning track for the **Brown University Datathon 2019** that covered Machine Learning & Natural Language Processing
- Won *Best Folded Protein* after completing 7-day Protein Folding and Molecular Modelling Workshop
- Elected as the President of Kalam Centre - A non-profit organisation dedicated to educating underprivileged kids (2017-18)
- Presented Poster on "Predicting gene expression under Enviromental Stress in *O sativa* at the MidSURE presentation Session

Additional Achievements and Awards

- **Indian National Informatics Olympiad**: Programming Competition that leads to the National Training Camp
  - Top 3.8% (2014)
  - Top 19% (2015)
- **Indian National Linguistics Olympiad**: National Training Camp that selects the National Team
  - Silver Medal (2015)
  - Honorable Mention (2016)
- **Indian Philosophy Olympiad**: National Team Selection: Final Round. National Rank: 11 (*2016*)
- **Baltic Sea Philosophy Essay Competition (UNESCO)**: National Rank: 8. Essay Score: 7.78/10.00 (*2017*)
- **Technothlon National Aptitude and Logic Competition (2014)**: Organised by Indian Institute of Technology, Guwahati. National Rank: 3
- **IRIS (ISEF National Science Fair (2012))**: Outstanding Award for Overall Project in Mathematical Sciences with a Silver Medal