

## Table of Contents

R265942 - 2024 University Recruiting - Applied Machine Learning Summer Intern

Candidate Name	Candidate ID	Attachments
1. Zhe Fang	C1531382	
2. Nathalie Valenzuela	C1585370	
3. AGNI RAJINIKANTH	C1585330	
4. Aditya Garg	C1482207	
5. Daniel Thompson	C714339	
6. Ravikiran Sriram	C1583930	
7. Julian Hunter	C1583842	
8. Palak Agarwal	C1583527	
9. Dhruv Thota	C1336566	
10. Yuxi Guo	C1582860	
11. Sai Chandra Pandraju	C1553288	
12. Amirsalar Bagheri	C1581930	
13. Pouria Mohammadi	C1581771	
14. Prateek Gautam	C1581537	
15. Eugene Han	C1574614	
16. Shuliang Yu	C1581050	
17. Ravikiran Sriram Sriram	C1580940	
18. Andre van de Ven	C1545157	
19. Sivaiah Naidu Yerraganti	C1574378	
20. Ethan Marshall	C1576201	
21. Maneth Balasooriya	C1580661	
22. Christine Park	C1580231	
23. Trisha Karani	C1504116	
24. pradeep Bolleddu	C1580167	
25. Qijia HE	C1580131	
26. Jessica Coulson	C1441390	
27. Jaden Reeves	C1579930	
28. Darshilkumar Dhameliya	C1579899	
29. Daniel Herbert	C1578132	
30. Luan Lam	C1048828	
31. Tom Earnest	C1579312	
32. Debanjan Goswami	C837419	
33. Jessica Mayer	C1567360	
34. Banmeet Kaur	C1578375	
35. Neel Patel	C1481900	
36. Can Li	C846721	
37. Austin Rush	C1533246	
38. Jonathan Wan	C1576112	
39. Elizabeth Tsang	C1470475	
40. Jordan Jomsky	C1575473	
41. Hang Yu	C1271373	
42. Roshan Patel	C1282981	

Candidate Name	Candidate ID	Attachments
43. DOLLY RIJWANI	C1575780	
44. Tanay Dixit	C1575419	
45. Ross Chu	C1575311	
46. Karthikeya Telapalli	C1573912	
47. Jaemin Yang	C1573861	
48. Xiangxi Tian	C1573601	
49. Jesse Na	C1275282	
50. Benjamin Zhao	C1572878	
51. Atharss Prasath	C1508113	
52. Sai Akhilesh Veldi	C1572226	
53. Kourosh Teimouri Baghaei	C1572112	
54. Yousuf Qari	C1136377	
55. Preston DeLeo	C1359884	
56. Yi Zhang	C1473398	
57. Wenjin Zhang	C1571791	
58. Jesse Graham	C1540588	
59. Hira Anees Awan	C1513014	
60. Maria Arce	C1571269	
61. Fengshuo Song	C1570167	
62. Charles Torres	C1569871	
63. Vaishnavi Meka	C1435935	
64. Owen Grimm	C1528800	
65. Tianmin Zhao	C1457812	
66. Meet Panjwani	C917598	
67. Lynne Auger	C1564059	
68. Jiayu Li	C1569365	
69. Moriyah Schick	C774920	
70. Xinwei Huang	C1568793	
71. Malini Shivaram	C1568804	
72. Shuliang Yu	C1563323	
73. Tharun Pasham	C1552839	
74. Mohammad Beheshti	C1071889	
75. Adel Abdalla	C1567264	
76. Aman Bhutani	C1497994	
77. Eli Edme	C1566702	
78. Bhavik Agarwal	C1566618	
79. Alessandro Hizdri	C1531683	
80. Marielys Reyes Castillo	C1564864	
81. Sean Neuman	C1564080	
82. Amogh Patankar	C1474522	
83. Sean O'Leary	C1563459	
84. Zubair Lakhia	C499095	
85. Christina Chandra	C1554835	
86. Santosh Tiwari	C1557319	
87. Hansoo Chang	C1481380	
88. jamal ansary	C1561806	

Candidate Name	Candidate ID	Attachments
89. Sonya Jin	C1560964	
90. Veronica Jenkins	C1560412	
91. Mason Kim	C1465930	
92. Yunqing Qiu	C1559752	
93. Nithin Parthasarathy	C1559761	
94. Shubham Gurav	C1237172	
95. Akshaj Kumar	C1559517	
96. Daniel Clepper	C1553302	
97. Charles Plonski	C1031349	
98. Gideon Popoola	C1559380	
99. ABDUL MATIN	C1559147	
100. Aderinsola Osinowo	C1263910	

# Zhe (Gigi) Fang

10 Akron St, Cambridge, MA 02138 | (617) 642-5698 | [gigifang925@gmail.com](mailto:gigifang925@gmail.com)

## EDUCATION

<b>Harvard University</b> , Cambridge, MA, USA	Sept 2019 – Expected 05/25
<i>Ph.D. in Epidemiology &amp; Biostatistics, Data Science</i> (secondary field)	Cumulative GPA: 3.99/4.0
<i>M.S. in Biostatistics</i>	Cumulative GPA: 4.0/4.0
<i>M.S. in Epidemiology</i> (awarded in June 2021)	Cumulative GPA: 4.0/4.0
<b>Peking University</b> , Beijing, China	Sept 2014 – June 2019
<i>M.B.B.S.</i> (equivalent to M.D. in the US) & <i>Bachelor of Economics</i>	Major GPA: 3.98/4.0

- Relevant courses: Machine Learning (at MIT), Causal Inference, Data Science, Statistical Inference and Regression, Survival/Multilevel and Longitudinal/Bayesian Analysis, Natural Language Processing, Computer Vision (at MIT)

## SKILL

- **Statistics:** Probabilities, Statistical inference and modeling (e.g., GLM, Cox, GLMM, GEE, GAM), Multivariate analysis, Bayesian analysis, Time series
- **Causal inference:** Randomized experiment design, Confounding control, Propensity score, Causal interaction /mediation analysis, IPW, g-formula, Instrumental variable, Regression discontinuity, Difference in difference
- **Data science:** Data mining, Supervised and unsupervised machine learning, Deep learning, NLP, CV
- **Other:** Epidemiologic study design and analysis, Decision analysis, Medicine
- **Programming:** Python (3 yr), R (4 yr), SAS (7 yr), Stata, Unix/Linux, SQL, GitHub, LaTex
- **Libraries:** NumPy, pandas, Matplotlib, seaborn, scikit-learn, TensorFlow, SciPy, PyTorch
- **Cloud environment:** Google Colab, FAS Research Computing Cluster, Domino Data Lab

## PROFESSIONAL EXPERIENCE

**Scientist Intern in Statistics and Computational Methods**, *Moderna*, Cambridge, MA      June 2023 – Aug 2023

- Worked on reactogenicity analysis of mRNA vaccine products based on metadata across three clinical trials.
- Performed data wrangling and exploratory data analysis, employed generalized linear mixed effect models/ordinal regression/censored regression, and pioneered addressing sex bias in reporting adverse reactions.
- Visualized results with ‘emmeans’ R package and drew causal graphs to evaluate potential bias in effect estimate.
- Streamlined the pipeline in Jira and Bitbucket, and gave presentations to stakeholders at various technical depths.

**Graduate Researcher in Big Data and AI for Medicine**, *Harvard University*, Cambridge, MA    Mar 2020 – Present

- Utilized multidimensional data including questionnaires/medical records/biospecimen in large population cohorts (N ~ 200,000) ranging over 30 years, processed into longitudinal databases, and performed survival analysis to investigate the effects of diet, metabolic biomarkers, and autoimmune diseases on cancer etiology.
- Developed an EHR-based AI risk assessment tool using machine learning techniques (e.g., LSTM, transformer model) to predict risk of colorectal cancer over time among patients with history of polypectomy in colonoscopy.
- Yielded 5 first-author publications and 4 coauthored publications in peer-reviewed journals.

**Research Assistant in Health Informatics**, *Usher Institute, University of Edinburgh*, UK      Sept 2018 – Dec 2018

- Built search strings for systematic literature search in 4 databases and extracted contents from eligible articles.
- Performed multilevel meta-regression to forecast the global prevalence of atherosclerosis in 2020 and random-effect meta-analysis to assess risk factors of atherosclerosis.
- Yielded 1 second-author publication that is acknowledged as a top 1% highly cited paper in 2020.

**Digital Health Co-op**, *the George Institute of Global Health*, Beijing, China      Dec 2017 – May 2018

- Led a six-member team for the project ‘Validation of Three Electronic Blood Pressure Monitoring Devices’.
- Worked cross-functionally by collaborating with software engineers and physicians on the App development to facilitate participant recruitment and data entry, trained and coached, and supervised fieldwork for data gathering.
- Conducted data analysis & data visualization on the validity comparisons between three blood pressure monitoring devices, wrote reports, and presented project results to stakeholders.

**Undergraduate Researcher in Causal Inference**, *Peking University*, Beijing, China      Jan 2017 – May 2019

- Conducted Mendelian randomization studies to infer the causal relationships between (1) 62 genome-wide factors and chronic kidney disease; (2) dairy intake, cardiometabolic traits, and body composition, performed statistical models on data from genome-wide consortia, assessed robustness of instrumental variable assumption.
- Designed regression discontinuity, a quasi-experimental approach, to examine the causal effect of fetal exposure to famine on adult obesity using data from China Biobank, published the paper, and developed a practical guide.

## PROJECT

### *Sentiment Analysis of Google Reviews for Tinder Application (Harvard CS109 Data Science)*

- Leveraged data science and NLP techniques such as data preprocessing/emoji representation/tokenization/word embedding, built pipeline for training and evaluating classification algorithms (*RNN*, *bidirectional LSTM*, *BERT*, *RoBERTa*) to classify sentiments expressed in their reviews (N=533,018).
- The best model (*RoBERTa*) achieved a test accuracy of 0.94.

### *Machine Learning for Suggesting Personalized Intervention for Obesity (MIT 6.793 Machine Learning in Healthcare)*

- Selected obesity-relevant phenotypic, genotypic, metabolic, and medical history & lifestyle features based on random forest and expert knowledge; reduced data dimension with Variational Auto-Encoder.
- Clustered patients in reduced dimension with Gaussian Mixture Models and K-Means, enhanced interpretability with optimal classification tree showing how different clusters can be explained using selected features.
- Mapped cluster to BMI-GWAS significant genetic variants and gene targets, nominated potential drugs.

### *Deep Learning in Brain Tumor MR Image Classification (MIT 6.8300 Advances in Computer Vision)*

- Acquired human brain MRI images from figshare, SARTAJ, and Br35H, did data augmentation, deployed deep convolutional neural network with transfer learning (*VGG*, *ResNet*, and *GoogleNet*).
- Evaluated model performance with various metrics including accuracy, F1 score, and confusion matrix.
- The best model (*VGG16*) achieved a test accuracy of 0.97.

## **ADDITIONAL EXPERIENCE**

<b>Academic Journal Reviewer</b> , <i>IJE/EJE/IJC/BMC Cancer/Scientific Reports</i>	Jan 2021 – Present
• Provided feedback on scientific novelty and significance, study design, statistical methods, and result interpretation of the research article submitted, suggested improvements, and recommended decisions.	
<b>Teaching Fellow</b> , <i>Harvard T.H. Chan School of Public Health</i> , Boston, USA	June 2021 – Aug 2022
• Served as teaching fellow for graduate courses <i>Design and Conduct of Trials in Preventive Medicine</i> , <i>Introductory and Advanced Statistics for Medical Research</i> , and won the Distinction in Teaching Award.	
<b>Intern Doctor</b> , <i>Beijing Shijitan Hospital</i> , Beijing, China	Feb 2017 – Feb 2018
• Did clinical rotations, spearheaded a six-member team assessing the medical expenditure of inpatients with heart diseases after Beijing Medicine Reform, and won 2 <sup>nd</sup> place in the University Entrepreneurial Competition.	

## **PUBLICATION**

1. Fang, Z., ..., & Song, M. Risk prediction models for colorectal cancer: evaluating the discrimination due to added biomarkers. *International Journal of Cancer*. 2021
2. Knudsen, M., ..., Fang, Z., ... & Song, M. Development and validation of a risk prediction model for post-polypectomy colorectal cancer in the USA: a prospective cohort study. *eClinicalMedicine*. 2023
3. Song, P., Fang, Z., ... & Rudan, I. Global and regional prevalence, burden and risk factors for carotid atherosclerosis: a systematic review and modeling analysis. *Lancet Global Health*. 2020
4. Fang, Z., ..., & Giovannucci, E. Association of nut consumption with risk of total and five specific cancers: Evidence from three large prospective cohort studies. *American Journal of Clinical Nutrition*. 2021
5. Fang, Z., ..., & Giovannucci, E. The role of Mendelian randomization studies in deciphering the causal effect of obesity on cancer. *Journal of National Cancer Institute*. 2021
6. Fang, Z., ..., & Song, M. Serum lipid profiles and risk of colorectal cancer: A prospective cohort study in the UK Biobank. *British Journal of Cancer*. 2020
7. Fang, Z., ..., & Tang, K. Association between fetal exposure to famine and anthropometric measures in adulthood: A regression discontinuity approach. *Obesity*. 2020
8. Fang, Z., ..., & Tang, K. The patterns and social determinants of breastfeeding in 12 selected regions in China: A population cross-sectional study. *Journal of Human Lactation*. 2019
9. Huang, T., Heianza, Y., Sun, D., Ma, W., Fang, Z., ... & Qi, L. Interaction of diet/lifestyle intervention and TCF7L2 genotype in glycemic control and adiposity among overweight or obese adults: big data from seven randomized controlled trials worldwide. *Health Data Science*. 2021

\*The full list of 19 publications can be found on my [Google Scholar](#)

## **AWARD & HONOR**

- Harvey Finberg Fellowship (2023), Finalist in Three-minute Thesis Competition of Nutrition 2023 Conference (2023), Certificate of Distinction in Teaching at Harvard (2022), Most Innovative Master Thesis at Harvard Chan (2021), Underrepresented Student Minority Engagement Award (2021), Brian and Heidi MacMahon Education Award (2021 & 2023)
- Graduation with Highest Distinction in Beijing (2019), Most Outstanding Student Award in Beijing (2018), National Scholarship (2018 & 2016), Merit Student Award at Peking University (2018 & 2016), Highest Scholarship for Medical Student at Peking University (2017), Excellence in Leadership Award (2016), Academic Excellence Award (2015)

# Nathalie Valenzuela



(209) 683 - 8318



nathalievalenzuela20@gmail.com

## EDUCATION



California State University, Fresno | Fresno, CA | **(Expected: Dec. 2023)**

### Masters of Science in Mathematics

Area of Research: The Role of California Fires in Predicting Valley Fever

California State Polytechnic University, Humboldt | Arcata, CA | **(May 2021)**

### Bachelors of Arts in Mathematics

## WORK EXPERIENCE



### TEACHING ASSISTANT

**Fresno State | Fresno, CA | August 2021 – December 2023**

- Courses: MATH 11- Elementary Statistics, MATH 45- What is Math?, Math 3- College Algebra
- Made assignments for students, prepared course materials, administered examinations, assessed student performance, including determining course grades. Attended meetings scheduled by the faculty for purposes of briefing, coordination, and discussion of course content and materials

### RESEARCH ASSISTANT

**Data Science Institute Summer Lab | Chicago, IL | Summer 2022**

- Explained technical concepts to nontechnical clients.
- Conducted analysis on the racial bias of traffic stops using the veil of darkness theory and other analysis models, and produced data visualization and maps with Python

### CAMP COUNSELOR

**Camp Military Teen Adventure Camp | Darrington, WA | Summer 2021**

- Developed and facilitated training for adult volunteers and camper orientation. Assisted in the set-up, maintenance, and tear-down of camp activities and spaces

### RECREATION INTERN

**Glacier Peak Institute | Darrington, WA | Summer 2021**

- Planned, coordinated, and facilitated youth programs focused on STEM, outdoor recreation, and environmental education. Assisted in critiquing and revising existing programming, problem-solved unexpected changes in activity plans, camper, and/ participation as a team.

### ADMINISTRATIVE ASSISTANT

**Arcata Forest Products | Arcata, CA | March 2020 - April 2021**

- Worked in Excel and Word Doc, Maintained current and accurate customer files and information, Organized and maintained file system, files correspondence, retrieved data and other records upon request, entered Arcata Forest Products sales orders in Lumber Track Inventory Sales System and filed sales orders after shipping.

## SKILLS



- Proficient in Python, R
- Data Exploration and Analysis
- Time management and organization
- Problem- solving skills

# AGNI RAJINIKANTH

(201) 824-6469 · agni.the.great@berkeley.edu · <https://r-agni.github.io/Portfolios/>

## EDUCATION & COURSEWORK

### **UC Berkeley - Management, Entrepreneurship, & Technology Program**

Dual Degree: B.S. Electrical Engineering & Computer Science/B.S. Business, SEED Honors Scholar, Google Scholar

**SOPHOMORE**

GPA: 4.0

*Blockchain Researcher at Center for Responsible Decentralized Intelligence under Prof. Dawn Song*

*Co-Founder@Berkeley STEP Startup Incubator, Space Technology@Cal(Electrical Engineering) Time II Mission, Blue Origin Strategy Consultant*

## EXPERIENCE

### **Ecotone Renewables[\$8 million startup], Pittsburgh, PA/USA**

**JUN 2021-PRESENT**

#### **Project Manager, Automation Leader, Shareholder (2021 – present)**

- Facilitated Google Design Springs, OKR Planning as lead of 4 CMU CS/ENG. students to convert food waste to renewable energy/fertilizer Met with 40+ stakeholders in the Northeast
- Built ML, Computer Vision, CNN, YOLOv4 model: 93% accuracy, Robotic Arm Manipulator, Hydraulic System For waste segmentation

### **NASA Katherine Johnson IV&V Facility, Fairmont, VA/USA**

**MAY 2022-AUGUST 2022**

#### **Machine/Quantum Learning Intern**

- Developed ML Computer Vision Analysis (CNN U-NET) Classification for Moon Rovers to navigate lunar surfaces autonomously. Technologies: Qiskit, OpenCV, Keras, TensorFlow, scikit-learn
- Deployed software testing and assurance testing for Machine Learning and Quantum solutions including Unit, Regression, Integration, Minimum Functionality Testing

### **HP Tech Ventures, Palo Alto, CA/USA**

**DEC 2022-FEB 2023**

#### **Corporate VC & Business Analytics**

- Conducted research on various aspects of potential startup investments, such as market size, growth rate, competitive landscape, product differentiation, and revenue model leveraging Pitchbook and Crunchbase
- Sourced deals and screened companies by performing due diligence on startup teams, traction, financials, and valuation deploying quantitative and qualitative methods, such as DCF, comparable analysis, SWOT analysis

### **Learnroll Immerse LLC, Somerset, NJ/USA**

**NOV 2023-JUNE 2023**

#### **AI/ML Data Engineering**

- Conducted 2-week scientific experiment alongside Prof. Tamorish Kole bringing low cost/affordable medical education through AR/VR
- Evaluated responses to medical emergencies through computer vision, image segmentation and 3d cameras

## LEADERSHIP & COMMUNITY IMPACT

### **United Nations and NGOs, NYC, NY/USA**

**SEPT 2021-SEPT 2023**

#### **Blockchain/ML Researcher&Developer, Youth Rep. to UN DGC**

- Piloted ML program for students in Indonesia in partnership with UN's Innovation Academy and Theta Pan Hub. Technologies: AI, Pandas, NumPy, Matplotlib
- Developed a Blockchain/NFT based platform for NGOs to obtain funding/grants to increase visibility in partnership with the Global NGO Executive Committee. Technologies: Ethereum, Blockchain, Smart Contracts

### **MEGA, NYC, NY/USA**

**SEPT 2021-PRESENT**

#### **President, Co-Founder**

- Lead a team of 30 from the USA, Canada, India, and China, hosted 3 events to promote UN sustainability goals: 500+ participants, 100+ projects, raised \$800,000+ supported by AOPS, GNEC, Resilience, Startup Brite, Twilio, Axure, Hyper Ex, Jet Brains, FISCA, DINO, Pukka Tech
- Currently building a streamlined challenge-based NGO/Startup recruiting pipeline solution

## AWARDS & ACHIEVEMENTS

- VigiVolc[Natural Disaster Detector using Infrasound ]: Business Innovator Award(2022), Patent#: 202221003798
- TSBand[Tourette Syndrome Smart Bracelet]: Publication in JMIR Neurotechnology First Author, Patent#: #63/378,863 Startup(Exited)
- New York Academy of Sciences: Young Scientist Award [Eutrophication with Dr. Ogugah, MedTech with Dr. Chellappan]
- Published Econometrics paper on Blockchain/Futures Markets as First Author, Mentored by Prof. Lincoln[Claremont McKenna]
- Horus[Personal Eye Care Assistant]: Anthony Parker Memorial Plaque: British Computing Society(2022)
- 1st Author Research on Beehive CCDs, Computer Vision, LoraWann at Liberty Science Center, Mentored by Prof. Martin[Rutgers-WINLAB]
- Harvard Spring Debate World Cup: Finalist(2022)
- American Computer Science League- Senior Division- All-Star Finalist(2021)
- Excellence in Quantum Machine Learning – IBM Qiskit

## ADDITIONAL INFORMATION

- Computer/Programming: Python, Java, C, Solidity, Html/CSS, JS, Dart/Flutter, TensorFlow, Qiskit, Pandas, NumPy, Matplotlib, Firebase
- Concepts: AI/ML, Blockchain, Quantum Computing, IOT, Robotics, Hardware, CAD, LoraWAN
- Soft Skills: Agile Scrum Management, Project Management, Spring Facilitation, Public Speaking, Marketing

# Aditya Garg

P: (408) 338-5920 | garkaditya@gmail.com | <https://www.linkedin.com/in/adigarg12> | <https://github.com/AdiG1123>

## EXPERIENCE

---

### MG Systems and Software LLC

*Data Science/Machine Learning Intern*

San Ramon, CA

Feb 2023 – Present

- Achieved a 40% reduction in advertising expenses by conducting in-depth market and customer analysis, optimizing product positioning, and implementing data-centric marketing strategies by mining through email, sales, and leads data.
- Engineered a Python-driven ETL pipeline (Pandas, pyodbc) to create an interactive sales conversion rate analysis dashboard in PowerBI. Identified process improvements, yielding a 10% increase in direct leads, equivalent to nearly \$100,000 in additional revenue, and informed future sales strategies.
- Streamlined the invoice billing system through automation, utilizing pywin32, Pymupdf, PyPDF2, and Pandas significantly decreasing invoice turnaround times and improving invoice validation processes.
- Delivered a complete ML solution with Docker, OpenCV, and TensorFlow for unstructured video data processing, including frame splitting, binary classification with MobileNetV2, and a Express.js API.

### Unchained Labs

*Research Associate*

Pleasanton, CA

May 2022 – Feb 2023

- Designed, proposed, and conducted experiments for feasibility and validation of hardware and software for lipid nano particle buffer exchange in tight timelines ranging from 2 weeks to one month.
- Delivered data-centric insights at over 35 meetings to scientists, engineers, and management.
- Created Python scripts to automate data retrieval from logs, reducing turnover of data analysis by 48 hours.

### Meissa Vaccines

*Virology Research Assistant II*

Redwood City, CA

Dec 2019 – April 2022

- Utilized Python to independently automate the reading of our clinical assay MN assay data that increased the efficiency of clinical assay output by 1-2 days.
- Created a simple LIMS system with Amazon Lightsail with MEAN stack incorporating storage for sequences and storing related sequence metadata.

## PERSONAL PROJECTS

---

### MCIT Community Hub (<https://github.com/TDeV4/nextjs-boilerplate>)

- Engineered a unified community-centric platform facilitating seamless navigation of vital information for users.
- Led the technical design and implementation with input of two team members utilizing Git and Trello to track project milestones allowing us to code the application within 2 months.
- Designed and normalized database schema using ER-diagram and wrote complex SQL queries for REST API.

## SKILLS

---

Languages: Python, Java, SQL, C, JavaScript, R

Cloud: AWS RDS, Amazon Sagemaker, AWS EMR, AWS S3, Amazon Athena, AWS Glue

Libraries: Pandas, Dask, Pyspark, MLlib, OpenCV, Scikit-learn, SciPy, NumPy

Tools: Git, Docker, Jupyter, Microsoft Office, Postman, PowerBI, Apache Spark, Google Colab

Frameworks: Express.js, React.js, TensorFlow/Keras, PyTorch

Domain Knowledge: Virology, Immunology, Cell Biology, Genomics, Molecular Biology, Physiology

## EDUCATION

---

### University of Pennsylvania (UPenn)

Dual Degree *Master of Science in Data Science | Master in Computer and Info Tech*

Aug 2022 – Dec 2024

- 3.88 GPA

Relevant Coursework: Big Data Analytics, Database and Information Systems, Intro to ML, Stats for DS

### University of California, Santa Barbara (UCSB)

*Bachelor of Science in Cell and Developmental Biology*

July 2016 – Dec 2019

# DANIEL THOMPSON

Phone: (305) 904-3314 • Address: 60 Morningside Drive, New York, NY 10027  
Email: [daniel.thompson1406@gmail.com](mailto:daniel.thompson1406@gmail.com) • LinkedIn: [linkedin.com/in/daniel-thompson1406](https://linkedin.com/in/daniel-thompson1406)

## EDUCATION

### Columbia University | New York, NY

Bachelor of Science in Chemical Engineering, Minor in Computer Science:

Expected May 2025

- **GPA:** 3.74/4.00
- **Major Concentration:** Biotechnology & Biopharmaceuticals, and Data & Computational Science
- **Relevant Coursework:** Machine Learning for Biomolecular and Cellular Applications, Organic Chemistry I & II w/Lab, Differential Equations, Calculus-Based Statistics and Probability, Discrete Mathematics, Applied Python, Thermodynamics, Transport Phenomena

## PROFESSIONAL EXPERIENCE

### Merck & Co. | Rahway, NJ

Manufacturing Strategy, Business Development & Alliances Intern:

June 2023 - August 2023

- Led a cross-functional IT team in creating an impactful dashboard using Power BI, effectively visualizing the division portfolio's impact across 50 manufacturing sites. Led to streamlined project selection processes, optimizing resource allocation and decision-making.
- Enhanced efficiency in the Leadership Development Program Alumni Database by automating data retrieval through Power Apps, Power Automate, Power BI, and SQL. Led to improved program analysis, heightened impact evaluation, and increased alumni retention.
- Developed an internal presentation deck outlining the functional domains within the Manufacturing Strategy, Business Development & Alliances department, facilitating seamless alignment with enterprise strategies and fostering collaboration with key stakeholders.
- Spearheaded the implementation of diversity, equity, and inclusion initiatives for the leadership board of the MSBDA department.

### Rovis Research Group | New York, NY

Research Assistant:

October 2022 - June 2023

- Conducted research on metalloenzymes by doing 10 hours of lab work a week, running metalloenzymes through GC analysis, analyzing data using computational techniques, and compiling final results for presentations and journal submissions.
- Analyzed Rh(III)-cofactors that enable C-H activation chemistry in aqueous environments

### NASA L'SPACE Mission Concept Academy | Tempe, AZ

Deputy Project Manager:

January 2022 - May 2022

- Managed and collaborated with a cross-functional team of engineers and scientists to develop a Preliminary Design Review for a theoretical lunar mission project
- Conducted research to identify optimal lunar landing sites using JMARS, a planetary Geospatial Information System software

### Laboratory for Complex Materials and Devices | Los Angeles, CA

Research Assistant:

September 2021 - May 2022

- Developed ionic semiconductors through pulsed laser deposition and implemented them on devices used for memory and imaging
- Researched chalcogenide perovskites and analyzed their efficacy as ionic semiconductors

## LEADERSHIP & PROFESSIONAL DEVELOPMENT

### Brazilian Society at Columbia | New York, NY

Treasurer:

Sep 2023 - Present

- Maintained accurate and up-to-date financial records for the Brazilian Society at Columbia
- Managed a \$4,000 budget utilized to organize events for 175+ members

### D.E. Shaw | New York, NY

Latitude Fellow:

May 2023 - Present

- Selected as 1 of 40 rising juniors nationally to attend DESCO's Latitude Fellowship for minority groups.
- Investigated unique trading strategies and techniques to leverage market fluctuations.

### Management Leadership for Tomorrow | Washington, DC

Career Prep Fellow:

December 2022 - Present

- Selected among thousands of high-achieving applicants to participate in this Career Preparation program focused on professional and leadership development
- Completed business case studies and assignments to hone analytical, quantitative, and technical skills

### Society of Hispanic Professional Engineers | Los Angeles, CA

Director of the New Student Advisory Board:

September 2021 - May 2022

- Onboarded 20+ students from Hispanic and Latinx backgrounds by organizing and conducting social events
- Attended leadership development modules and implemented skills learned within those modules to lead new members.

## SKILLS

- **Technical Skills:** Power BI, Power Apps, Power Automate, Excel, PowerPoint, Python, C++, Java, SQL, HTML, MatLab
- **Languages:** English (Fluent), Portuguese (Fluent), Spanish (Advanced)

# RAVIKIRAN SRIRAM

Address Jersey City, New Jersey Phone (425) 624-1436 E-mail [sriramravikiran7@gmail.com](mailto:sriramravikiran7@gmail.com) [LinkedIn](#)

## Education

<b>Stevens Institute of Technology</b>	Hoboken, NJ
Master of Science: <b>Business Intelligence and Analytics</b>	Sep 2023- Dec 2024
Relevant coursework: Optimization and Process Analytics, Multivariate Analytics, Business Intelligence and Data Integration	
<b>Padre Conceicao College of Engineering- Goa University</b>	Goa, India
Bachelor of Engineering: <b>Computer Engineering</b> (GPA:3.5/4)	Jul 2016 – Aug 2020
Relevant coursework: Artificial Intelligence, DBMS, Data Mining, Data Structures, Algorithms, Operations Research, Cloud Computing	

## Work Experience

<b>InvestMan Financial Services</b>	Panaji, Goa
<b>Analyst Intern</b>	Jan 2023- May 2023
<ul style="list-style-type: none"><li>Analyzed customer goals and predicted an asset investment plan with 85% accuracy using Pandas and Numpy, resulting in a 20% growth in client base</li><li>Spearheaded a data cleaning and validation initiative, resulting in a 30% improvement in data accuracy and a 20% reduction in processing time.</li><li>Conducted in-depth data mining and trend analysis to provide actionable insights for strategic decision-making, contributing to a 15% increase in quarterly revenue</li></ul>	
<b>Claysys Technologies</b>	Panaji, Goa
<b>Software Engineer (L2)</b>	Feb 2022- Jun 2022
<ul style="list-style-type: none"><li>Designed web applications and cross-platform mobile applications using ASP.NET MVC and Xamarin according to client briefs and modified designs to meet changes in client specifications</li><li>Led a team of 5 developers to create applications and perform validation and verification testing and collaborated with internal teams to fix and improve products</li><li>Built stored procedures and triggers to perform automated rules, updating to related tables in SQL server along with creating tables, views, index and relations</li></ul>	
<b>Persistent Systems</b>	Verna, Goa
<b>Software Engineer</b>	Oct 2020- Jul 2021
<ul style="list-style-type: none"><li>Developed web applications utilizing HTML5, CSS, Javascript and Angular JS and integrated RESTful APIs</li><li>Implemented CSS techniques and leveraged insights from Google Analytics to enhance the loading speed of high-traffic pages, resulting in a 20% improvement in user experience</li><li>Gathered and defined customer requirements and collaborated with cross-functional teams to ensure precise specifications for project plans</li></ul>	

## Skills

<b>Technical skills</b>	Python, R, Java, C, C++, C#, HTML5, CSS, Javascript
<b>Python Packages</b>	Numpy, Pandas, Scikit- Learn, NLTK, TensorFlow, Regex, Keras
<b>Databases</b>	SQL, MySQL, Oracle, MongoDB
<b>Software Tools</b>	Microsoft Suite(Excel, Word, PowerPoint), Microsoft Visio, Jira, Snowflake, Tableau, Hadoop, Alteryx, Erwin
<b>Data Visualization</b>	Matplotlib, ggplot2, Seaborn, Power BI, Tableau
<b>Certifications</b>	Data Analysis with Pandas and Python, Google Data Analytics ,Big Data Analytics using Hadoop and Spark

## Academic Projects

<b>Neural Network for Palm Tree Disease Detection</b>	Jun 2019- Oct 2020
<ul style="list-style-type: none"><li>Created a neural network utilizing deep learning and machine learning methodologies utilizing Google Colab to train a model using an 800-image dataset for early detection and classification of palm tree diseases, achieving a robust accuracy rate of 98%.</li></ul>	
<b>IOT Based Smart City Bus Transport</b>	Aug 2019 – Oct 2020
<ul style="list-style-type: none"><li>Developed an Android app for Goa's public transit, proposed to the Government to increase usage and monitor commuters. Linked users to Google Cloud via Raspberry Pi, optimizing routes based on real-time traffic using a modified Dijkstra's Algorithm.</li></ul>	
<b>Airline Delays And Cancellation Analysis</b>	Jan 2022- Mar 2022
<ul style="list-style-type: none"><li>Utilized Excel and Tableau to analyze a dataset from January 2015 to March 2015, containing over 1 million records, to create a dashboard highlighting the best and worst airline performance.</li></ul>	
<b>New York Airbnb Rental Analysis</b>	May 2022- Aug 2022
<ul style="list-style-type: none"><li>The analysis of a dataset of New York-based Airbnb listings from 2011 to 2019 resulted in the creation of a dashboard with information on the locations with the highest rental frequency, their impact on rental costs, and a chart of host acquisition by region.</li></ul>	

# Julian Hunter

[ikrhunt72@gmail.com](mailto:ikrhunt72@gmail.com) | [www.linkedin.com/in/julian-hunter-b9487b246/](https://www.linkedin.com/in/julian-hunter-b9487b246/)

2995 Egan Dr, Memphis, TN 38115 / (901)572-3227

**Profile:** Current College Sophomore. Civic minded, dependable, and mature. Currently seeking internship opportunities.

**Education:** White Station High School, Memphis TN. 2022. Expected College Graduation Date: May 2026  
Current GPA: 3.72 / Weighted GPA: 4.241

Attending NCAT, Studying Mechanical Engineering

Dual Enrollment Completed: Engineering I; Engineering Prob. Solving

## Employment

- *Leidos: Power Delivery Intern (Summer 2023)*
  - Develop plans and schematics for the transmission poles vital to the transportation of energy across the country; do so within a typical work environment
- *YMCA: Lifeguard Intern (Summer 2022)*
  - Assign wristbands to and inform members/nonmembers before entering the pool, learn from supervisors of the tasks lifeguards must perform, supervise and take care of pool
- *MPLOY: City of Memphis (Summer 2019; 2020;2021)*
  - Utilized OSHA standards to standardize good workplace ethics and encourage positive behavior
  - Learn jazz music from instructors and perform for public events (2019)

## Skills:

- |                       |                               |                     |                    |                 |
|-----------------------|-------------------------------|---------------------|--------------------|-----------------|
| • Strong Leadership   | • Creativity                  | • Critical Thinking | • Team Cooperation | • Adaptable     |
| • Respectful          | • Staying Calm Under Pressure |                     | • Tenacity         | • BOUD Software |
| • SolidWorks Software |                               |                     |                    |                 |

## Leadership:

- Marching Band Sub-Section Leader: (2021-2022)
- Senior Patrol Leader- Boy Scouts: (2021-2022)
- Eagle Scout Project (2021)

---

## Extracurricular Activities

### Community based:

- NAACP ACT-SO Competition (2021)
- Memphis Challenge Ambassador (2019- 2022)
- Memphis Jazz Workshop (2019)
- Boy Scouts of America: (2009-2022)

### Eagle Scout

- National Society of Black Engineers Jr. (2013- 2022)
- Make a Splash Swim Team (Splash Midsouth) (2013-2018)

### School based:

- All West Tennessee Band (2022)
- Spruce Up Sparta (2020- 2022)
- Dungeons & Dragons Club (2021-2022)
- Marching Band (2018- 2022)
- Chess Club (2018-19)
- Swim Team (2018)

---

## Community Service:

- Eagle Project- Bridge at T.O Fuller State Park (2021)
- Spruce Up Sparta (2020-2022)
- Mobile Pantry Food Drive (2020)
- Volunteer Odyssey (2020)
- Voter Registration Encouragement (2020)

- Humane Society (2020)
  - Exceptional Foundation (2019)(2022)
  - Clean Memphis (2019)
  - March of Dimes (2019)
  - Girl Scout Cookie Distribution volunteer (2018-19)
  - Boy Scouts (Multiple projects 2009- present)

---

## Honors and Awards:

National Honors Society (4/12/22)

Eagle Scout Rank (1/11/22)

Alpha Lambda Delta Honors Recipient

National Recognition Scholar

# Palak Agarwal

San Jose, CA • +1 (415) 510 9765 • [pakags.15@gmail.com](mailto:pakags.15@gmail.com) • [www.linkedin.com/in/palakagarwal124/](https://www.linkedin.com/in/palakagarwal124/)

## SUMMARY

Graduate student with prior experience of **4+ years** as a **Data Engineer** in building & scaling data-intensive distributed applications. Strong CS fundamentals with proficiency in architecting & implementing robust real-time and batch data processing pipelines to enable analytics.

## EDUCATION

### San Jose State University, San Jose, California

Aug 2022 - May 2024

#### Master of Science (MS) in Computer Software Engineering (Big Data)

CGPA: 3.95/4.0

**Courses:** Big Data Algorithms, Big Data Engineering and Analysis, Engineering Analysis, System Engineering, Database Systems, Topics in Emerging Technologies (CPS with AI), Foundations for Decision and Data Science (R programming)

### Jaypee Institute of Information Technology, India

2014 - 2019

#### Bachelor of Technology + Master of Technology in Computer Science and Engineering

CGPA: 7.9/10

**Courses:** Soft Computing, Deep Learning, Information Integration and Visualization, Data Structures and Algorithms for Big Data, Machine Learning and Data Mining, Cloud Based Enterprise Applications, Social Media Analysis, OOP, Database Systems and Web, Data Structures

## TECHNICAL SKILLS

**Languages:** SQL, C++, C, Php, Scala (Beginner), Python, Java (Beginner), Pyspark, R programming and HTML 5

**Databases & Data Warehouse:** MySQL, Postgres, MongoDB, Redshift, RDBMS, Cassandra, Object Relational Mapping, NoSQL

**Big Data / ETL Tools:** Apache Spark, Hadoop, Flink (Beginner), Kafka, Hive, Airflow, Map Reduce, Abinitio, CDM (Campaign), Snowflake

**AWS (Cloud Platform):** EMR, Kinesis, IManaged Airflow, EKS, DynamoDB, Rekognition, Lambda, EC2, Step Function, Pinpoint, RDS, API Gateway, S3, SNS, Cloud Watch, Step Function, Athena, Terraform (IaC), AWS CLI, Kubernetes (Beginner)

**Miscellaneous:** Data Structures, Algorithms, Power BI, Tableau, MS Excel, Curl, Problem Solving, Distributed Systems, Large Scale Systems, Docker, Cloud Services, Data Analytics, Machine Learning, Data Mining, Git hub, Jupyter, Linux, Putty, Db Visualizer, Agile, Spark streaming, Looker, Qlik, Jenkins, CI/CD Pipelines

## RELEVANT WORK EXPERIENCE

### Autodesk, San Francisco | Data Engineer Intern

May 2023 - Aug 2023

- Automated transformation and export of excel files from SharePoint to pyspark reducing the data processing time to 25% for dashboard update.
- Created temporary views by performing UPSERT transactions in ETL tables for incremental load in hive to upgrade revenue tracking dashboard.
- Performed comparative analysis in Snowflake and Hive for multiple tables in excel to develop data pipelines which reduced human efforts by 40%.

### The Data Team, India | Data Engineer

Mar 2021 - Apr 2023

- Developed the feature to automatically on-board the new attributes in Cadenz User Profiling and the subsequent trigger notifications using SNS.
- Created AWS Data Ingestion Pipelines using pagination to query DynamoDB using AWS lambda for serving the subsequent API calls.
- Implemented the Flink framework for calculating the attributes for upto 180 days by maintaining their states on different realtime Kafka Streams.
- Re-orchestrated and migrated the entire system architecture from airflow to AWS services which reduced map-reduce processing time by 60%.

### FICO, India | Solution Integration Engineer

Feb 2019 - Mar 2021

- End-to-end responsibility of delivering the brand campaign workflows using segmentation techniques to improve quality of data pipelines.
- Automated the Data Extraction and Suppression via effective ETL scripts in Abinitio for Data formatting and Transforming in Redshift.
- Built the employee survey platform for the quarterly measure of employee happiness index.

## ACADEMIC PROJECTS

### Identification of Item Ratings using Collaborative Filtering

Aug 2022 - Dec 2022

- Implemented the user - user collaborative filtering technique using pyspark to identify the item ratings for test dataset without using ALS library.
- Compared accuracy for SVD and LSH methods using RMSE score in HPC with 6 nodes in parallel and reduced processing time to few minutes.

### Spotify Songs Analyzer

Aug 2022 - Dec 2022

- Executed filtering and modelling techniques to identify user's most liked songs for top artists over time using big data in Cassandra.
- Numerous visualizations on tableau and R facilitated predictions for top genre and the span for a song to remain in the top list.

## PUBLICATIONS

- P. Agarwal, M. Sharma and G. Kaur, "Content-based Classification Approach for Video-Spam Identification", Springer 17th International Conference on Intelligent Systems Design and Applications (ISDA 2017). [doi: [https://doi.org/10.1007/978-3-319-76348-4\\_23](https://doi.org/10.1007/978-3-319-76348-4_23)]
- P. Agarwal, M. Sharma and S. Chandra, "Comparison of Machine Learning Approaches in the Prediction of Terrorist Attacks", IEEE 12<sup>th</sup> International Conference on Contemporary Computing (IC3 2019). [doi: <https://doi.org/10.1109/IC3.2019.8844904>]

## CERTIFICATIONS

- Certified** AWS Cloud Practitioner, by Amazon Web Services (AWS), Issued on May 11, 2020, Validation Number S1FPN32KGMQEQQ6KG
- Certified** Azure Fundamentals, by Microsoft, Issued on June 23, 2021, Certification Number H868-0499

# Dhruv S. Thota

(732) 666-4163 · 101 Susan Circle, North Wales, PA · dst5222@psu.edu · [Linkedin](#)

---

## EXPERIENCE

<b>Research Assistant, Teacher Shortages Project, PA Alliance for STEM Education</b>	<b>Summer 2022</b>
<ul style="list-style-type: none"><li><i>Conducted research for Dr. John Weathers on teacher turnover and shortages through literature reviews and National Center of Educational Statistics' School and Staffing Survey data analysis.</i></li><li><i>Used Google Scholar to search for relevant empirical literature and create brief summaries of key findings</i></li><li><i>Prepared and cleaned data for statistical analysis in SPSS.</i></li><li><i>Conducted basic statistical analysis to examine relationships and differences across teacher and school characteristics.</i></li></ul>	
<b>Board Secretary, Future Business Leaders of America,</b>	<b>September 2021 - June 2022</b>
<ul style="list-style-type: none"><li><i>Managed organization meetings, records, and communication for a 100 member society</i></li></ul>	
<b>National Honors Society</b>	<b>2021- 2022</b>
<b>Key Club</b>	<b>2021- 2022</b>
<b>Syracuse Biomedical Engineering Summer Program for HS Students</b>	<b>Summer 2021</b>
<ul style="list-style-type: none"><li><i>Program that introduces students to the biomedical world as I created an app that helps people who think they might have migraines accurately determine if they have one or not</i></li></ul>	
<b>Delta Epsilon Vice President/ Phi Coordinator, German Club</b>	<b>September 2020 – June 2022</b>
<ul style="list-style-type: none"><li><i>As VP, I made sure everyone in the cabinet was on track and planned an honors event as Coordinator.</i></li></ul>	
<b>State Records Recorder, Skippack Pharmacy, Lansdale, PA and Skippack, PA</b>	<b>Feb 2021 - Present</b>
<ul style="list-style-type: none"><li><i>Volunteered 30+ hours in a month period and helped put information of vaccine in state files</i></li></ul>	
<b>Food and Service Champion, Taco Bell</b>	<b>July 2020 – February 2021</b>
<b>Consulting Internship for client (through Value-N-Action Consulting in Toronto)</b>	<b>June 2023 - Present</b>
<ul style="list-style-type: none"><li><i>Project related to launch of a new innovation hub in Africa to serve startups and entrepreneurs in the developing world. Hoping to serve multiple industries such as Pharmaceutical, Oil, Agriculture, etc.</i></li></ul>	

## EDUCATION

<b>Pennsylvania State University, Bachelors of Science</b>	<b>Anticipated Graduation May 2026</b>
<ul style="list-style-type: none"><li><i>Cumulative GPA: 3.15</i></li><li><i>Student of the College of Engineering in the Biomedical Engineering Program</i></li><li><i>Significant Coursework - Chemistry 1 &amp; 2, Chemistry 1 &amp; 2 Lab, Math 140 (Calculus and Analytic Geometry), Math 141 (Calculus 2), Engineering Design 100</i></li><li><i>Member of the PSU Chapter of the Biomedical Engineering Society and member of the Society of Industrial Biotechnology</i></li></ul>	
<b>North Penn High School, Lansdale, PA</b>	<b>Graduated June 2022</b>
<ul style="list-style-type: none"><li><i>Weighted GPA: 5.30</i></li><li><i>Significant Coursework - AP Economics, Biology, Chemistry, Algebra 2, Pre-Calculus, AP Computer Science Principles, AP Statistics, Physics, Principles of Engineering.</i></li></ul>	

## SKILLS

- I possess good analytical, communication and problem solving skills*
- I can smoothly conduct meetings and I have attained leadership skills as I have taken in many roles*
- I am experienced with the python language*
- I am experienced in SolidWorks which is CAD software where I can create 3D structures*
- I have received 7 certifications in Biotransformation from the SIB (Society of Industrial Biotechnology)*

## QUALIFICATIONS

A highly skilled self-motivated professional with one year of experience in **data manipulation, quantitative modeling** and **statistical analysis** who is looking for a **data analytics intern** position

## CORE KNOWLEDGE & SKILLS

- Experience with data mining, statistical analysis, predictive modeling and machine learning techniques
- Extensive experience utilizing R, Python, SQL for data manipulation, modeling and visualization
- Strong logical assessment and interpretation skills
- Assume responsibility and take initiative
- Excellent communication and interpersonal skills
- Able to effectively straddle both technical and business realms

## PROJECTS

- Developed and honed professional skill set by undertaking class and personal projects      05/2022 - Current
  - Predicted probability of heart disease and discovered the main factors impacting the probability  
Conducted data exploration and preprocessed dataset using feature engineering, dimensionality reduction, and visualization techniques. Predicted probability of heart disease by using multiple machine learning techniques including Logistic Regression, Random Forest and Boosting in R. Identified the main factors impacting the probability by analyzing feature importance
  - Developed customer segmentations and profiled segments  
Preprocessed dataset and built customer segmentations based on customer shopping recency, frequency, and monetary value using unsupervised learning techniques in Python. Profiled each segment based on customer historical spend and demographic information. Created and delivered analytic insights and recommendations in presentations to the class
  - Performed data analysis on retail data  
Conducted data exploration on retail data of 100K+ records and identified strategies for the retailer to conduct comprehensive customer lifetime value analysis, enabling targeted marketing and retention strategies

## PROFESSIONAL EXPERIENCE

- Sr. Research Scientist, Chemily Glycoscience LLC, GA, USA      05/2019-02/2022
  - Led all aspects of product development lifecycle  
Partnered with Sales team and led chemical product development from end to end, delivered innovative products and met customer needs. Facilitated sales team to exceed the annual sales goal by 9%.
  - Conducted chemistry intensive research and make recommendations  
Conducted chemistry intensive research and applied knowledge of products for the completion and successful execution of projects. Recommended opportunities and avenues for process improvement
  - Developed a close relationship with internal and external clients  
Established and maintained effective working relationships with internal sales team and external clients to understand their needs. Delivered findings and recommendations in succinct and compelling presentations to both technical and non-technical audiences.

## EDUCATION

- Master of Science in Analytics, Georgia Institute of Technology, USA      05/2022 - Current
- Doctor of Science in Chemistry, Georgia State University, USA      08/2014 - 05/2019

## AWARDS

- 2016 International Carbohydrate Symposium, New Orleans, Honorable Mention Award

# Sai Chandra Pandraju

Boston, MA | +1 (857)-693-9901 | pandraju.s@northeastern.edu

[LinkedIn](#) | [GitHub](#) | [Portfolio](#)

Available: Summer 2024

## EDUCATION

**Northeastern University**, Boston, Massachusetts, USA

September 2022 - Present

**Khoury College of Computer Sciences**

Expected Graduation: December 2024

Master of Science in Computer Science

GPA: 3.83

Related Courses: Machine Learning, Database Management Systems, Programming Design Paradigm

**Jawaharlal Nehru Technological University**, Kakinada, India

April 2019

Bachelor of Technology in Electronics and Communication

Related Courses: Computer Programming in C, Data Structures, Probability & Statistics, Linear Algebra & Differential Calculus

## TECHNICAL SKILLS

**Languages** Python, SQL, Java, C, Typescript, HTML, CSS

**Libraries and Tools** Scikit-Learn, NumPy, Pandas, Keras, TensorFlow, PyTorch, Transformers, Ray, PySpark, DeepSpeed, Xgboost, NLTK, spaCy, SciPy, SHAP, Lime, LIT, Matplotlib, GIT, Docker

**Frameworks** Django, Flask, Spring Boot, Angular

**Cloud Services** AWS (Lambda, Sagemaker, EC2, S3, Auto scaling), Azure (AzureML), Google Cloud (Vertex AI, Cloud Storage, BigQuery, Compute Engine, Firebase)

**Certifications** TensorFlow certified developer, Mathematics for Machine Learning Specialization, Deep Learning Specialization

## PROFESSIONAL EXPERIENCE

**SoftInWay Inc.**, Burlington, Massachusetts, USA

June 2023 - Present

Machine Learning Intern

- Designed, developed, and deployed an **end-to-end grounded response chatbot** system for SoftInWay's wiki, resulting in a **reduction of tech support's efforts** in addressing client queries by **>70%**
- Leveraged **Airflow** to create an efficient data pipeline for scraping, cleaning, chunking, embedding, and storing embeddings in a vector store, **optimizing retrieval process**
- Retrieved and re-ranked the top 5 relevant wiki chunks per user query from vector store using **dot product similarity**
- Employed **dialog inpainting** technique to construct a dataset capturing conversational essence of user interactions for fine-tuning
- Fine-tuned** Large Language Models (Falcon-7B-instruct, Flan-alpaca-xl, LLaMA-2-7b-chat) by leveraging Fully Sharded Data Parallel (**FSDP**), **DeepSpeed** and **Ray Train** for optimized multi-GPU utilization
- Architected and created a relational database (**PostgreSQL**) to store user details, interactions and feedback
- Built **asynchronous RestAPIs**, integrated PostgreSQL interaction with Ray's ActorPool, **dockerized**, and efficiently **deployed** chatbot inference with **dynamic request batching** on multi-GPU using **Ray Serve**
- Connected **Grafana** and **Prometheus** with the Ray Serve dashboard, enabling visualization of various deployment metrics, including request count, request latency, load per replica, and resource utilization

**Infosys Ltd.**, Bengaluru, Karnataka, India

August 2019 - July 2022

Senior Software Engineer - Machine Learning

- Improved the **F1 score** of '*Biomedical Relation Extraction System*' by **15%** through the integration of transformer-based models
- Conducted Exploratory Data Analysis (**EDA**) and implemented **LDA** and **DBSCAN** to **cluster** the ticket database of over **9.5 million records** and incorporated solution into a real-time ticket allocation application
- Improved accuracy of the alert system by **20%** through the application of diverse **time-series techniques** on SAP HANA tables' record count trend
- Adopted a **Data-Centric approach**, utilizing **Stanford Lab's Snorkel AI slicing capabilities**, resulting in Infosys achieving a top **10** position (#**6**) in the **SuperGLUE** benchmark, a rigorous evaluation for **NLU** tasks
- Developed Python libraries for **autoML** training, **custom script** training, and **deployment** modules, utilizing **AutoGluon**, **SeldonIO's MLServer**, and **MLFlow**, as part of the SDK for Infosys' low-code AI platform

## RELEVANT PROJECTS

**Make More**, (Language Modelling)

May 2023 - June 2023

- Trained and compared diverse language models, including count-based, single-layer, multi-layer embedding-based, RNN, and transformer's decoder architectures, **all built from scratch in PyTorch**

**Few-Shot Entity Recognition**, (Multi-modal)

February 2023 - April 2023

- Constructed efficient **few-shot learners** for extracting **unseen entities** from receipt images, employing a **prototypical network** approach with pre-trained models (LayoutLM and LiLT) as feature extractors

# Amirsalar Bagheri

---

Manhattan, KS | Phone: 773-280-1561 | Email: [amirbg@ksu.edu](mailto:amirbg@ksu.edu) | [linkedin:Amirbg4](#) | [Github:amirbg4001](#)

## Education

### **Ph.D. of Chemical Engineering**

**Kansas State University | Manhattan, KS**

**Expected Graduation: December 2025**

- Machine learning-based estimation and optimal control of chemical and agri-food processing systems

### **Master of Science, Chemical Engineering**

**Kansas State University | Manhattan, KS**

**Expected Graduation: May 2024**

- Machine Learning-based Process Control and optimization.

### **Bachelor of Science, Chemical Engineering**

**Amirkabir University of Technology | Tehran, Iran**

**June 2020**

- Design and manufacture of a temperature control system for a batch reactor using Raspberry Pi

## Summary

Chemical Engineering Ph.D. student with 3+ years in Data Science, process modeling, and optimization. A dedicated team player, I merge data science with chemical engineering in my research. Skilled in deep learning for optimal control in chemical and food industries, proficient in TensorFlow and PyTorch. Seeking an internship in Summer 2024 to further develop my communication and teamwork abilities.

---

## Awards

**KSU Graduate Student Council (GSC) travel award**

**September 2023**

**KSU Graduate Research and State presentation award**

**October 2023**

## Presentations

### **Learning-Based Data Reconstruction and Predictive Modeling of an Ammonia Synthesis Process for State Estimation and Control Applications.**

**November 2023**

Amirsalar Bagheri, Thiago Oliveira Cabral, Davood B.Pourkargar. AIChE Annual meeting 2023, Orlando, FL

### **A Deep Learning-Based Model Reduction and Control of an Ammonia Synthesis Process (Contributed)**

Thiago Oliveira Cabral, **Amirsalar Bagheri**, Davood B.Pourkargar, AIChE Annual meeting 2022, Phoenix, AZ

## Relevant Experience

### **A Recurrent Neural Network (RNN) Simplification for Ammonia Reactor's Predictive Control.**

**July 2022**

- Served as a key member of a team project to resolve the computational bottleneck of the ammonia reactor model for optimization.
- Used TensorFlow to develop a Long short-term memory (LSTM) framework for the ammonia production reactor by receiving illuminating advice from the major advisor.
- For integrating the LSTM model with Model predictive control (MPC), communication between Python, MATLAB, and COMSOL Multiphysics was achieved due to effective teamwork.
- Results submitted to the American Control Conference (ACC 2024)

### **A physics-informed machine learning approach for predicting soil water content**

**May 2022**

- Dedicated member of an effective interdepartmental collaborative project comprised of 3 departments.

- Utilized Python and TensorFlow to develop a physics-informed machine learning strategy for teaching soil physics to neural networks with the help of collaborators from K-State College of Geology and College of Agriculture
- Results submitted to the American Control Conference (ACC 2024)

### **A Machine Learning Solution for plant-based meat food extrusion modeling**

**August 2022**

- Effectively collaborated with top researchers in the Food Science Department of Kansas State University
- As a team player, I used PyTorch for data-driven approaches such as RF, SVM, and FNNs to model a food extrusion system based on a real plant-based meat food extrusion process.
- Designed an Image processing system with OpenCV and Canny Edge detection algorithm for binary classification of plant-based meat.
- Designed an RNN-based network with TensorFlow to capture extruder dynamics based on real extrusion data.

### **Distributed Machine Learning for Faster Complex Integrated Process System Modeling.**

**July 2023**

- Efficiently communicated with teammates to perform relevant model reductions to build the dataset
- Pytorch with MPI library is used for training models on CPU nodes
- Explored federated learning solutions for accelerating and optimizing ML-based process modeling

## **Leadership Experience**

### **Kstate Iranian Graduate Students Association (IGSA)**

**January 2023**

- Organized and managed social, sports, and Persian cultural events.
- Worked with university officials to play a major role in helping financially troubled students.

## **Other Experience**

### **Core Multi-processor course**

**May 2021**

- Designed a software-based neural network accelerator in MATLAB for a jpg compression algorithm
- Analyzed CPU performance in terms of synchronous clocks for the neural network accelerator.

### **Computer Architecture**

**June 2021**

- Designed an MIPS machine with Verilog and implemented ALU operations
- Implemented pipelining strategy for parallel CPU computing (Multiple Instructions)

## **Skills**

### *Personal Skills*

Problem-solving, team player, critical thinking, effective communication

### *Programming Languages*

Python, PHP, C++, C#, HTML5 , CSS3, JavaScript, MATLAB,Verilog

### *Web Frameworks*

Laravel, ASP.NET and MVC5, Django

### *Database*

SQL, MySQL (DBMS), SQL Server (DBMS)

### *Operating systems*

Linux OS (Fedora/Ubuntu/Cent OS/Kali)

### *Machine Learning*

TensorFlow,Keras, Pytorch,Pytorch-lightening

### *Computer vision*

OpenCV

## **Certificates**

- LPIC 1 & 2** – Kahkeshan Noor Institute **August 2018**
- Developing web applications with ASP.NET and MVC5** – Tehran Institute of Technology **August 2017**
- Web development with .NET Framework** – Tehran Institute of Technology **May 2017**
- Programming in HTML5 with JavaScript and CSS3** – Tehran Institute of Technology **August 2016**
- C++ Programming** – Tehran Institute of Technology **June 2016**

## **References**

**Sajid Alavi** – Professor at Dept. of Grain Science and Industry - Kansas State University  
*Contact:* Waters Hall 51 Manhattan, KS 66506 – [salavi@ksu.edu](mailto:salavi@ksu.edu) - 785-410-9051

**Davood B.Pourkargar** – Assistant Professor at Dept. of Chemical Engineering - Kansas State University  
*Contact:* 2017 Durland Hall, 1701A Platt St., Manhattan, KS 66506 – [dbpourkargar@ksu.edu](mailto:dbpourkargar@ksu.edu) - 785-532-2625

# POURIA MOHAMMADI

## Summary

Ph.D. student in civil engineering pursuing a minor in a graduate certificate in data science, almost 2.5 years of experience in developing machine/deep learning algorithms. Saved millions of dollars in maintenance costs for more than 100,000 culverts in the State of Utah by delivering a regression model that predicts culvert conditions with 80% accuracy. The anticipated graduation date is 12/25.

## Skills

- 
- Python (Pandas, NumPy, SciPy, Matplotlib, OpenCV, PyTorch, Scikit-Learn, and Scikit-Image)
  - Data visualization and Data Wrangling
  - MATLAB, Fortran, LaTeX
  - SQL, Microsoft Access, Excel
  - Tableau, Microsoft Power BI

## Projects

<b>ESTIMATING UTAH'S CULVERT DETERIORATION CURVES</b> - Utah Department of Transportation's Project	August 2022
<ul style="list-style-type: none"><li>Utilized <b>Python</b> to analyze 2070 culvert inspection data</li><li>Implemented <b>Random Forest Regression</b> and <b>Support Vector Regression</b> algorithms to predict the condition of culverts</li><li>Aggregated and visualized the data by using <b>Pandas</b> and <b>Matplotlib</b> to compile a professional report</li></ul>	
<b>EVALUATING ML ALGORITHMS FOR AUTOMATED CULVERT INSPECTION</b> - Research Project	October 2023
<ul style="list-style-type: none"><li>Analyzing 2555 culvert inspection data with <b>Python (Scikit-learn)</b></li><li>Conducted a comprehensive evaluation of five ML algorithms (<b>Random Forest, SVM, K-NN, ANN, Decision Tree</b>)</li><li>Enhancing the assessment of culvert condition through increasing accuracy to 82%</li></ul>	
<b>ROAD PAVEMENT CRACK DETECTION</b> - Personal Project	December 2022
<ul style="list-style-type: none"><li>Utilized <b>PyTorch</b> to perform analysis on 6,004 images of US roads from the <b>RDD2022</b> dataset</li><li>Leveraged the <b>FasterRCNN</b> neural network architecture to detect and quantify 6 types of pavement cracks</li></ul>	
<b>MEASURING SEMANTIC SIMILARITY BETWEEN ARTICLES AND PROJECTS</b> - Personal Project	April 2023
<ul style="list-style-type: none"><li>Performed data cleaning and text analysis on the abstracts of articles and project reports in <b>Python</b></li><li>Using pre-trained <b>transformer-based</b> models (<b>Sentence-BERT</b>) to convert texts into embedding vectors</li><li>Apply <b>K-means clustering, DBSCAN, and BERTopic</b> algorithms to cluster articles in recent Civil Engineering literature</li><li>Computed the <b>cosine similarity</b> between embedded vectors of technical reports and academic articles</li></ul>	
<b>IMPLEMENTING FEDERATED LEARNING FOR CONDITION PREDICTION</b> - Research Project	September 2023
<ul style="list-style-type: none"><li>Developed a <b>federated-learning-based</b> model on 8,240 culvert inspection data from six states' inventories</li><li>Utilized <b>FLOWER</b> and <b>PyTorch</b> libraries in <b>Python</b> for the development of federated learning and compared its results with centralized learning</li><li>Fine-tuned <b>ANN</b> model performance to achieve 80.4% accuracy using the <b>Federated Averaging</b> algorithm</li></ul>	

## Work Experience

<b>GRADUATE RESEARCH ASSISTANT</b> - University of Utah - Salt Lake City, UT	August 2021 - Present
<ul style="list-style-type: none"><li>Working on various data science projects, including transportation asset management with <b>tabular</b> and <b>image</b> data</li><li>Utilizing <b>optimization</b> algorithms (e.g. <b>Swarm Intelligence</b>) to fine-tune <b>XGBoost</b> models in predicting culvert condition ratings</li><li>Managed two undergraduate students in data collection sessions (project management skills)</li><li>Writing a proposal to <b>automate</b> culvert inspection procedure using <b>deep learning</b> and <b>computer vision</b>, which is funded by the Utah Department of Transportation (\$60,000)</li></ul>	

## Education

<b>DOCTOR OF PHILOSOPHY IN CIVIL ENGINEERING</b> - University of Utah - Salt Lake City, UT	August 2021 - Present
<b>MASTER OF SCIENCE IN CONSTRUCTION MANAGEMENT</b> - University of Tehran - Tehran, Iran	Sept 2017 - Aug 2020
<b>BACHELOR OF SCIENCE IN CIVIL ENGINEERING</b> - University of Tehran - Tehran, Iran	Sep 2013 - Aug 2017

## Profile

Computer science educator and a startup co-founder well-versed in translating data into actionable insights. Passionately driven to utilize my teaching experience, analytical background, and programming expertise to tackle business challenges.

## Education

**Georgia Institute of Technology, Atlanta, GA** January 2023 - Present  
**Master of Science in Analytics**

- Computational Data Analytics specialization

**Ramapo College of New Jersey, Mahwah, NJ** Aug 2015 - May 2019

**Bachelor of Science in Mathematics** Minors in Computer Science and Economics

- Recipient of the most prestigious merit-based Presidential Scholarship (valued at \$120k+ over 4 years)
- Dean's List (5/8 semesters of full-time enrollment)

## Selected Projects/Participations:

- *Goal Scoring Analysis of Top Tier Soccer Clubs*: Implemented a multiple regression model to predict the goals scored by Europe's 114 soccer teams in the 2020-2021 season.
- *Interdisciplinary Contest in Modeling (ICM)* under the mentorship of Prof. Dr. Amanda Beecher. Project: Out of Gas and Driving on E (analysis of EV and combustion engine cars)
- *Medical Data Visualizer* that explored the relationship between cardiac disease, body measurements, blood markers, and lifestyle choices.
- *Page View Time Series Visualizer*: Data visualizations to help understand the patterns in user visits to websites and identify/predict monthly and yearly growth.
- *Data-driven Teaching*: Analyze student responses across multiple assessments (MCQs and Free Responses) to inform instructional decisions in and out of the classroom.
- Secretary of //hackRamapo club and President of Students For Liberty

## Certification and Technical Skills

- IBM Cognitive Class - Data Science Certificate
- DBMS and Data Analysis: MySQL, R, Python 3, MS Excel (forecasting, VBA, and Macros)

## Experience

**Zest Wealth Management** **Co-founder and Data Analyst** November 2021 - Present  
Kathmandu, Nepal, and Lyndhurst, New Jersey, USA

- Use advanced analytics techniques to provide personalized wealth management services to clients to help them achieve their financial goals. Analyzed data from 3000+ sources to guide marketing and product strategies leading to 3x traffic on the website and a 55% sales lift.
- Lead a team of three full-time and two part-time employees across two continents to achieve a long-term vision of maximizing returns and minimizing risks in the Nepal Stock Exchange (NEPSE).
- Demonstrated ability to analyze market trends, assess risk, and make informed investment decisions. For the FY ending in 2022, Zest achieved a 25% YoY net growth in its portfolio.
- Proven success in building and maintaining client relationships, resulting in 90%+ client retention.
- Portfolio under management in February 2023: NPR. 12 million.

**Digital Pioneers Academy** **High School Computer Science Teacher** September 2019 - Present  
Washington, DC

- Courses taught: AP Computer Science Principles, AP Computer Science A, AP Precalculus, and Geometry (small group tutoring/intervention).
- Successful track record of implementing and evaluating STEM programs through analytics, resulting in improved student achievement and engagement. Students achieved a 10% growth in PARCC and 85% of students taking the AP CSP test passed it (exceeding the national average of 67%).
- Expert in designing and delivering curricula aligned to state and national standards supported by a strong understanding of current educational trends and best practices in STEM education.
- Adept in using technology to enhance and support students' in-person and online learning.

# Eugene Han

eugeneh2@illinois.edu · (201) 841-2870 · eugenehan.github.io

## EDUCATION

---

### University of Illinois at Urbana-Champaign, Urbana-Champaign, IL

Ph.D. Statistics

Aug 2021 - May 2025

M.S. Statistics | GPA: 3.90/4.00

Aug 2019 - May 2021

### Carnegie Mellon University, Pittsburgh, PA

B.S. Mathematical Sciences with MCS Research Honors | GPA: 3.20/4.00

Aug 2015 - May 2019

Additional Major in Statistics, Minor in Computer Science

## PROFESSIONAL EXPERIENCE

---

### Sandia National Laboratories, Albuquerque, NM | Remote

*Intern Year Round - R&D Grad*

Aug 2022 - Sep 2023

- Built anomaly detection methods for acoustic signals and images in Python (TensorFlow)
- Advised summer interns on their corresponding projects
- Continued previous battery work conducted as a summer intern

### *Math & Analytics Graduate Intern*

May 2022 - Aug 2022

- Constructed failure forecasting models in R using Isolation Forests for lithium-ion batteries
- Published an internal manuscript detailing the methodology and results

### Locus Analytics, New York, NY

Jun 2018 - Aug 2018

#### *Data Analytics Intern*

- Developed classification models in Python to classify job postings to the firm's proprietary classification system
- Analyzed economic complexity of geographically proximate communities using clustering algorithms in Python

### Opticlose, New York, NY

Sep 2014 - Aug 2015

#### *Data Science Intern*

- Built models in R to predict the success of sales closure given tabular sales data

## RESEARCH EXPERIENCE

---

- [1] Policy Learning with Continuous Actions Under Unmeasured Confounding, *In Progress*.
- [2] E. Han and R. Zhu, Modeling and Visualizing Compositional Data with the Fisher-Bingham Distribution. *Joint Statistical Meetings (JSM)*, 2023. **Oral Presentation**.
- [3] E. Han and D. Offner, Linear  $d$ -polychromatic  $Q_{d-1}$ -colorings of the Hypercube, *Graphs and Combinatorics*, 34 (2018) 791-801.

## HONORS & AWARDS

---

University of Illinois at Urbana-Champaign Block Grant Fellowship

May 2022

Dean's List (High Honors)

Spring 2019 (2018)

Top 20% on the Putnam Mathematical Competition (score: 11, 13)

Dec 2017, 2018

Recipient of the Harris Award at HackCMU

Sep 2015

## PROGRAMMING SKILLS

---

R, Python, C, Java, Git

# Shuliang Yu

120 Meyer Rd, 1-505,  
Buffalo, NY. 14226

716-730-1468  
[mogystat@outlook.com](mailto:mogystat@outlook.com)

## Objective

Looking for an intern position of Statistician/Data scientist where I can apply my skills on statistical modeling and computation incorporating my strong domain knowledge on neurobiology and health related fields.

## Education

### State University of New York at Buffalo

Aug 2021 - present

- Ph.D candidate in Biostatistics
- MA in Biostatistics
- Relevant coursework: Limit Theories, Theory of Statistical Inference, Advanced Theories of Linear Models, Advanced Statistical Computing, Categorical Data Analysis, Longitudinal Data Analysis, Survival Analysis

### Rhodes College

Aug 2016 - Aug 2020

- BS in Neuroscience with Rhodes Award for four years
- Relevant coursework: Python, Java, C++, Molecular Neuroscience, Cell Biology, Neuroendocrinology, Movement Neuroscience, Organic Chemistry I & II, Analytical Chemistry, Physiology & Anatomy I & II, Biochemistry

## Statistical Research

### Applications of Mixture Models With Frequentist and Bayesian Approaches

May 2023 - present

Advisor: Dr. Saptarshi Charkraborty, Department of Biostatistics, UB

1. Accomplished an ECM algorithm and researched on the related mathematical properties
2. Extended the method to a multi-level Bayesian structures cooperating with Gibbs sampler
3. Designed and constructed an iterative structure of the ECM algorithm on R
4. Conducted simulation study to compare the method with competitors

### Applications of Supervised and Unsupervised Data Mining

1. Built tree-based classification (Random forest) and regression models (CART) to capture primary features in large-scale data sets
2. Made rules and predictions with structural learning algorithms and Apriori algorithms
3. Applied clustering algorithms such as K-means, Principal Component Analysis and Hierarchical Clustering to segment data and uncover hidden patterns using a Self-Organizing Map
4. Prevented overfitting in models with Lasso and Ridge regularization

### Computation and Simulation Study on Generalized Linear Model and Random Effect Model

1. Practiced on *tidyverse* for data cleansing
2. Fitted models with GLM methods (Linear regression, Logistic regression, Log-linear regression) and interpreted under different scenarios and different types of data.
3. Applied Bootstrap for further statistical inferences such as Goodness-of-fit test, hypothesis testing on parameters, interval estimates etc.
4. Created insightful visualizations by *ggplot2* and *plotly* and wrote scientific report in Markdown

### Association between Anxiety and Cognitive Impairment in Aging Mice

1. Handled challenges with longitudinal data, missing data, data entry errors, noisy data, multicollinearity and outliers
2. Implemented data cleansing, transformation and visualization to verify model assumptions and facilitate interpretation
3. Performed regression analysis and tested on correlation by Pearson's Chi-Square test to demonstrate association between anxiety and cognitive impairment

### The Best Linear Combinations of Multiple Biomarkers

1. Proposed a shrinkage method to calculate AUC of the best linear combinations of multiple biomarkers

2. Derived the large sample properties of the proposed Stein's type estimate
3. Designed a bootstrap algorithm for computing the p-values and confidence intervals
4. Conducted Monte Carlo simulations under different settings to demonstrate utility

## **Neurobiological Research**

---

### **Ferroptosis Regulating MiRNAs**

1. Cultured, froze and cryopreserved glioblastoma cell cultures
2. Conducted literature research on ferroptosis-related genes and pathways
3. Conducted Co-Immunoprecipitation, Western Blot, Gel Electrophoresis, and qPCR to compare between selected miRNA-regulated cells and control cells

### **Epigenetic Markers of Fetal Alcohol Spectrum Disorder**

1. Conducted literature research on the relationship between genetics and ethanol-related phenotypes in both the developing embryo and the adult animals
2. Dissected mice and used strategies such as PCR and LC to identify the critical genes and epigenetic markers, e.g. DNA methylations, in mediating a phenotype and determine its role
3. Drafted for the textbook *Springer Clinical Medicine - Neurodevelopmental Pediatrics: Genetic and Environmental Influences*

### **Effect of Corticosterone on Oxytocin Expression in the Hypothalamic Paraventricular Nucleus of a Chronic Stress Model**

1. Designed and applied experimental methods including Behavioral Interaction, Elevated Zero Maze, Open Field Chamber, Light-Dark Apparatus, Novel Object Test, Hyponeophagia test (HNP), Enzyme-Linked Immunoassay (ELISA)
2. Conducted data analysis using Mann-Whitney U test, Principal Component Analysis and Spearman's rho, etc.

### **Network Synchronization Across the Longitudinal Axis of the Developing Rat Hippocampus**

1. Measured extracellular recordings in slice models of epilepsy and analyzed network level activation
2. Simulated and analyzed effects of this model across the longitudinal axis of the hippocampus throughout development using NEURON and ImageJ

## **Publication**

---

Book chapter

- [1] Baker, J.A., **Yu, S.**, Scott, M.T., Hamre, K.M. (2023). Genetics and Epigenetics of FASD. Neurodevelopmental Pediatrics. *Springer, Cham.* [https://doi.org/10.1007/978-3-031-20792-1\\_27](https://doi.org/10.1007/978-3-031-20792-1_27)

## **Extracurricular**

---

□ Co-founder & Secretary of UB Buffalo ASA student chapter	2023-Present
□ Teaching Assistant for STA119: Statistical Methods, UB	2022-Present
□ Vice President of the Graduate Student Association, UB	2023-Present
□ Secretary of the Graduate Student Association, UB	2022-2023
□ Departmental Assistant at the Math, Chemistry and Political Science Departments	2016-2020
□ Piano Accompanist at Rhodes Orchestra, Harp Ensemble, and Chamber Trio	2016-2020
□ Resident Assistant at Rhodes College Residence Life, Rhodes College	2017-2019
□ Library Help Desk Worker, Rhodes College	Fall 2016

## **Technical Skills**

---

- Proficient in R, SAS, SPSS, Excel, LaTEX, Minitab, Rmarkdown, Quarto
- Experienced in Python, Java, C++, NEURON, FIJI ImageJ, MATLAB

# RAVIKIRAN SRIRAM

Address Jersey City, New Jersey Phone (425) 624-1436 E-mail [sriramravikiran7@gmail.com](mailto:sriramravikiran7@gmail.com) [LinkedIn](#)

## Education

<b>Stevens Institute of Technology</b>	Hoboken, NJ
Master of Science: <b>Business Intelligence and Analytics</b>	Sep 2023- Dec 2024
Relevant coursework: Optimization and Process Analytics, Multivariate Analytics, Business Intelligence and Data Integration	
<b>Padre Conceicao College of Engineering- Goa University</b>	Goa, India
Bachelor of Engineering: <b>Computer Engineering</b> (GPA:3.5/4)	Jul 2016 – Aug 2020
Relevant coursework: Artificial Intelligence, DBMS, Data Mining, Data Structures, Algorithms, Operations Research, Cloud Computing	

## Work Experience

<b>InvestMan Financial Services</b>	Panaji, Goa
<b>Analyst Intern</b>	Jan 2023- May 2023
<ul style="list-style-type: none"><li>Analyzed customer goals and predicted an asset investment plan with 85% accuracy using Pandas and Numpy, resulting in a 20% growth in client base</li><li>Spearheaded a data cleaning and validation initiative, resulting in a 30% improvement in data accuracy and a 20% reduction in processing time.</li><li>Conducted in-depth data mining and trend analysis to provide actionable insights for strategic decision-making, contributing to a 15% increase in quarterly revenue</li></ul>	
<b>Claysys Technologies</b>	Panaji, Goa
<b>Software Engineer (L2)</b>	Feb 2022- Jun 2022
<ul style="list-style-type: none"><li>Designed web applications and cross-platform mobile applications using ASP.NET MVC and Xamarin according to client briefs and modified designs to meet changes in client specifications</li><li>Led a team of 5 developers to create applications and perform validation and verification testing and collaborated with internal teams to fix and improve products</li><li>Built stored procedures and triggers to perform automated rules, updating to related tables in SQL server along with creating tables, views, index and relations</li></ul>	
<b>Persistent Systems</b>	Verna, Goa
<b>Software Engineer</b>	Oct 2020- Jul 2021
<ul style="list-style-type: none"><li>Developed web applications utilizing HTML5, CSS, Javascript and Angular JS and integrated RESTful APIs</li><li>Implemented CSS techniques and leveraged insights from Google Analytics to enhance the loading speed of high-traffic pages, resulting in a 20% improvement in user experience</li><li>Gathered and defined customer requirements and collaborated with cross-functional teams to ensure precise specifications for project plans</li></ul>	

## Skills

<b>Technical skills</b>	Python, R, Java, C, C++, C#, HTML5, CSS, Javascript
<b>Python Packages</b>	Numpy, Pandas, Scikit- Learn, NLTK, TensorFlow, Regex, Keras
<b>Databases</b>	SQL, MySQL, Oracle, MongoDB
<b>Software Tools</b>	Microsoft Suite(Excel, Word, PowerPoint), Microsoft Visio, Jira, Snowflake, Tableau, Hadoop, Alteryx, Erwin
<b>Data Visualization</b>	Matplotlib, ggplot2, Seaborn, Power BI, Tableau
<b>Certifications</b>	Data Analysis with Pandas and Python, Google Data Analytics ,Big Data Analytics using Hadoop and Spark

## Academic Projects

<b>Neural Network for Palm Tree Disease Detection</b>	Jun 2019- Oct 2020
<ul style="list-style-type: none"><li>Created a neural network utilizing deep learning and machine learning methodologies utilizing Google Colab to train a model using an 800-image dataset for early detection and classification of palm tree diseases, achieving a robust accuracy rate of 98%.</li></ul>	
<b>IOT Based Smart City Bus Transport</b>	Aug 2019 – Oct 2020
<ul style="list-style-type: none"><li>Developed an Android app for Goa's public transit, proposed to the Government to increase usage and monitor commuters. Linked users to Google Cloud via Raspberry Pi, optimizing routes based on real-time traffic using a modified Dijkstra's Algorithm.</li></ul>	
<b>Airline Delays And Cancellation Analysis</b>	Jan 2022- Mar 2022
<ul style="list-style-type: none"><li>Utilized Excel and Tableau to analyze a dataset from January 2015 to March 2015, containing over 1 million records, to create a dashboard highlighting the best and worst airline performance.</li></ul>	
<b>New York Airbnb Rental Analysis</b>	May 2022- Aug 2022
<ul style="list-style-type: none"><li>The analysis of a dataset of New York-based Airbnb listings from 2011 to 2019 resulted in the creation of a dashboard with information on the locations with the highest rental frequency, their impact on rental costs, and a chart of host acquisition by region.</li></ul>	

# Andre van de Ven

Philadelphia, PA | (480)-208-8823 | andrev@sas.upenn.edu

## EDUCATION

**University of Pennsylvania** | Philadelphia, PA

May 2026

*Candidate for Bachelor of Arts*

*Majors:* Biology, Computer Science

*GPA:* 4.0

**Relevant Coursework:** Programming Languages and Techniques, Data Structures and Algorithms, Big Data Analytics, Mathematical Foundations of Computer Science, Multivariable Calculus, Molecular Biology, General Chemistry, Neuroscience, Accounting

**Northern Valley Regional High School** | Demarest, NJ

June 2022

*Honors/Awards:* Dean's List; High School Salutatorian; National Merit Scholarship; AP Scholar with Distinction; NHS Officer

## PROFESSIONAL EXPERIENCE

**University of Pennsylvania Medicine, Researcher** | Philadelphia, Pennsylvania

March 2023 - Present

- Conducting research on the adverse effects of extended bisphosphonate therapy in treating bone metabolism disorders.
- Analyze large datasets from the UK Biobank, a biomedical database containing over half a million patients using Python.
- Employ machine learning models (logistic regression) to predict and identify patterns related to adverse effects in patients undergoing extended bisphosphonate therapy.

**NewYork-Presbyterian Hospital, Intern** | New York, New York

May 2022 – June 2022

- Gained exposure to a variety of surgical procedures, including coronary artery bypass grafts, meningioma removals, amputations, and DaVinci laparoscopic surgeries, expanding my understanding of surgical techniques and patient care.
- Engaged directly with dozens of surgeons, actively asking questions during surgeries to deepen my knowledge of surgical procedures, standard treatments, and intraoperative decision-making.
- Led a team of four other high school interns to fulfill staff requests for essential medical supplies and equipment.

**The Park Medical Group, Office Assistant** | Tenafly, NJ

January 2022 – May 2022

- Streamlined patient experiences by maintaining an accurate log for efficient scheduling and minimizing wait times, while also facilitating communication through the management of phone calls and appointment confirmations.
- Supported the medical team by taking vital signs such as pulse, blood oxygen saturation, and body temperature.

## LEADERSHIP & OTHER EXPERIENCE

**Divers at Penn, Treasurer** | Philadelphia, PA

April 2022 – Present

- Co-founded and manage finances for Divers at Penn, a club that unites students passionate about scuba diving/marine life.
- Raised over \$1000 to support the club's activities, including meetings and dive trips.

**Penn Spark, Social Chair/Developer** | Philadelphia, PA

September 2022 – Present

- Acquired comprehensive skills in full-stack web development through an intensive 10-week coding bootcamp, mastering technologies including React JSX, Django, Python, HTML, and CSS.
- Collaborated on the development of a user-friendly app to facilitate the creation of study groups on campus.
- As Social Chair, organize and lead engaging events to foster camaraderie and encourage the exchange of ideas.

**Penn Neuroscience Society, Research Committee** | Philadelphia, PA

September 2022 – Present

- Engaged in critical analysis and discussion of the latest advancements in neuroscience literature.
- Developed and delivered comprehensive presentations on cutting-edge developments such as Brain-Computer Interfaces (BCIs) and advancements in pharmaceutical research for neurodegenerative disorders.

**DECA Business Club, President** | Demarest, New Jersey

September 2018 – June 2022

- Advanced to the International DECA conference in both sophomore and junior years.
- Nominated to attend DECA's International Career Development Conference Leadership Academy
- Raised over \$5000 for local restaurants struggling during the onset of the pandemic; increased membership by >20%.

## SKILLS & INTERESTS

*Technical:* Java, Python (pandas, numpy), Ocaml, MATLAB, Javascript, SQL, Pytorch, Apache Spark, AWS, MLlib, SciKit-Learn

*Language:* Spanish (proficient), Dutch (proficient), Farsi (basic)

*Societies:* Wharton Investment and Trading Group, Phi Kappa Psi (Risk Chair)

*Interests:* Scuba diving and underwater photography. Winner of PADI club photography award. Find my photos at:

<https://www.andrevandeven.us/>

# Sivaiah Naidu

Buffalo, NY | [LinkedIn](#) | +1 716 416 0747 | [ysivaiahnaidu@gmail.com](mailto:ysivaiahnaidu@gmail.com)

## PROFESSIONAL SUMMARY

Candidate for data science intern with a master's degree in data science and applications from SUNY Buffalo. Extensive experience creating and testing data pipelines with cloud services and big data technologies such as AWS, GCP, Apache Airflow, and Apache Beam. Data processing and analysis skills in Python, Java, SQL, Hadoop, Hive, and Pyspark. Interested in applying my data science skills to real-world problems and learning from industry leaders.

## EXPERIENCE

### EXL Services Pvt. Ltd

Bangalore, India

Senior Consultant

July 2021 – July 2023

- Working for leading NBCU-Peacock multinational telecommunications conglomerate on a popular streaming service as a Data Engineer working on Apache Airflow and Kubernetes with Helm Chart deployed on Google cloud platform.
- As owner of Airflow deployment which works as a bridge for ML and other teams of the organization. I work on keeping the airflow cluster and underlying Kubernetes deployment up and running without issues.
- Developing and testing new Airflow pipelines as per the requirements from several end users. Orchestrating various GCP services like GCS, Big query.

### BizAcuity Solutions Pvt. Ltd

Hyderabad, India

Data Engineer

March 2020 – July 2021

- Working closely with business for developing Pyspark jobs using AWS Glue; developing the packages to extract data from Hive in EMR and transforming aggregate data into summary tables into Redshift
- Exploring various AWS services like EC2 instances, EMR clusters, Lambda functions, Glue, Step functions, Event bridge SNS services and many more to optimize the Data flow and as well as improving the development workflow.

### Hadoop-Hive Developer

September 2019 – March 2020

- As part of an internal initiative to explore various database performance at different scales using TPC Benchmark DS, I installed Hive and Hadoop on an AWS multi-node cluster and optimized its performance and resource scalability.
- Conducting Power and Throughput test replicating production level query execution scenarios which will help to analyze and implement various optimizations to the current configuration and Performing Impact Analysis of the configuration.

### ETL Developer

July 2018 – September 2019

- Working closely with business for developing ETL packages using ODI; developing the packages / interfaces to extract data from source and transforming aggregate data into summary tables.
- Developing and testing extraction, transformation, and load (ETL) processes; maintaining SQL scripts and complex queries for analysis and extraction.

## ACCOMPLISHMENTS & LEADERSHIP EXPERIENCE

### Secretary of Green Club, IIT BHU, July 2015 – March 2016

Varanasi, India

- Managing and organizing club events on and off campus that will drive students to be aware of various environmental issues and educate them on their responsibilities in this regard.

### IIT - JEE, Entrance examination, April 2014

- Secured a rank of 1836 out of 1.4 million in IIT - JEE, a highly competitive joint entrance exam for admission into elite Colleges under Indian Institute of Technology.

## EDUCATION

**State University of New York at Buffalo (SUNY-Buffalo)**

**Buffalo, NY, USA**

Masters of Professional Studies in Data Science and Applications, December 2024

**Indian Institute of Technology (IIT-BHU)**

**Varanasi, India**

*Bachelor of Technology:* Chemical Engineering, July 2018

**Narayana Junior College**

**Vijayawada, India**

Board of Intermediate Education Andhra Pradesh, March 2014

## SKILLS

**Certifications:** Microsoft Certified Solutions Associate SQL 2016 Database Administration Charter Member, Data Engineering, Big Data and Machine Learning on GCP Specialization on Coursera.

**Core Competencies:** Data Analysis, Cloud Services, Data warehousing, BI reporting and ETL processes.

**Programming Skills:** Python, Java, Shell Scripting, SQL, Hadoop, Hive, PySpark, AWS, GCP, Apache Airflow, and Apache Beam

# Ethan Marshall

<p><b>Contact</b></p> <p>908.202.0288 emarsha3@cougarnet.uh.edu Github: ethanmarshallanalytics</p>	<p><b>Objective</b></p> <p>Looking for a Data Science internship where I can apply the knowledge and skills that I have gained in the classroom and in prior internships to further develop my practical experience.</p>
<p><b>Education</b></p> <p>University of Houston Houston, TX M.S. Statistics and Data Science Anticipated Graduation: Aug 2024</p> <p>Drake University Des Moines, IA B.S. Data Analytics B.S. Mathematics Magna Cum Laude May 2023</p>	<p><b>Experience</b></p> <p><i>May 2022 – August 2022</i> Business Intelligence/Data &amp; Analytics Intern • Casey's</p> <ul style="list-style-type: none"><li>Worked in an Agile team, participating in retrospectives, sprint planning, and story completion (requirement gathering, project management tools, and report writing).</li><li>Worked and optimized reporting views for ad-hoc data pulls using SQL Server.</li><li>Built and refined metrics, attributes, reports and dashboards using MicroStrategy applications.</li><li>Lead a team of interns to developing and presenting a business case for a new marketing campaign.</li></ul>
<p><b>Technical Skills</b></p> <p>Machine Learning R Python SAS SQL &amp; T-SQL Tableau AWS Data Visualization Artificial Neural Networks Artificial Intelligence MicroStrategy Mathematics Jira Agile Project Management</p>	<p><i>May 2021 – August 2022</i> Fuel Pricing &amp; Supply and Transportation Analyst Intern • Casey's</p> <ul style="list-style-type: none"><li>Analyzed petroleum market trends and competitor movement to create pricing decisions at over 100 stores across Casey's footprint.</li><li>Developed new store performance analytics reports.</li><li>Created a program to project the demand of fuel by geographic market.</li><li>Constructed and presented a business model to add Deep Dish Pizza to Casey's menu.</li></ul> <p><i>May 2020 – May 2023</i> Admission Data Analyst &amp; Student Ambassador • Drake University</p> <ul style="list-style-type: none"><li>Analyzed population and census data to develop new marketing strategies for the school.</li><li>Evaluated Department of Education data to analyze data for new John D. Bright College.</li><li>Organized and delivered tours of Drake University to promote academics, athletics, and campus life with the goal of positioning prospective students to commit to Drake University.</li></ul>
<p><b>Honors &amp; Awards</b></p>	

---

Seven Dean's List Honors  
One President's List Honors  
Three Missouri Valley Conference Scholar-Athlete Team  
CSC Academic All-District  
USTFCCA All-Academic Award  
Eagle Scout Award

---

## Research Projects

Berkshire Hathaway Wind Turbine Modeling • May 2023  
Working with Berkshire Hathaway representatives, I partnered with fellow students to develop a model to predict when faults in wind turbines would occur. Presented our findings to Berkshire Hathaway to help sustain efficient wind turbine output to the greater Iowa area.

Projecting Whether an NCAA Men's Basketball Team Makes March Madness • November 2022

Using SAS and Python I created a Random Forest and Bernoulli Distribution model to predict which teams would qualify for March Madness. Results were statistically significant. Presented my research at the prestigious Midwest Sports Analytics Meeting and won best poster award.

---

## Leadership

Kappa Mu Epsilon (Mathematics Honor Society). Advisory Board  
Member for Drake University Bulldogs for Equity and Inclusion. Men's Track and Field Captain at The University of Houston. Men's Track and Field Captain at Drake University. Volunteer AAU Track and Field Coach.

---



# Maneth Balasooriya

260-580-6204 | [www.linkedin.com/in/maneth-balasooriya](https://www.linkedin.com/in/maneth-balasooriya) | <https://github.com/manethportfolio>  
[manethbalasooriya11@gmail.com](mailto:manethbalasooriya11@gmail.com)

## Tools and skills

**Languages :** Python, Java, C, Javascript

**Databases :** SQL, MongoDB, Postgre SQL, ETL and pipeline engineering, Redshift Clusters

**Frameworks/ML :** Node JS, Django, React, Tensorflow, Pytorch, Cuda and Jax

**Data Visualization :** Tableau, Numpy, PowerBI, Matplotlib, Pandas and Seaborn

**Tools :** AWS, Azure, Apache Spark, Apache Kafka, Git, Docker, scikit learn , REST API, HIVE, Snowflake

**Certificates:** Machine Learning Certificate by Deep Learning Labs, Data Warehouse fundamentals certificate

## Work experience

Purdue University library team

May 2021 - December 2022

Digital lab Assistant

- **Database Development and Library Management:** Orchestrated the development of a **PHP and MySQL-based database**, complete with **server setup and login integration**, laying the groundwork for a comprehensive library management system to efficiently catalog and access archive files from various sections.
- **3D Scanning and VR Guidance:** Instructed students and tutors in the use of 3D scanning technology and aided in integrating scanned objects into a virtual reality environment, enhancing educational experiences.
- **Digital Archive Creation:** Meticulously curated and created archive files combining text and image formats for seamless upload to the database, preserving historical and artistic content for future reference and research.

## Projects

[EDA - Real Estate project evaluation in NYC](#)

*Google Colab, Matplotlib, Seaborn, Pandas, Numpy*

- Conducted in-depth statistical analysis to create a comprehensive analytics report.
- Employed Matplotlib and Seaborn for data visualization, producing insightful charts and graphs.

[Data Visualization](#)

*Tableau*

- Leveraged Tableau to craft interactive and informative dashboards.
- Worked with Netflix and Airbnb datasets, transforming raw data into actionable insights for decision-makers.

[Stock Market Data pipeline](#)

*Python, Kafka, AWS S3, AWS Glue, AWS Athena, MySQL*

- Developed a robust data pipeline for real-time data ingestion.
- Utilized Kafka as an orchestrator and **AWS S3 buckets** for data storage.
- Implemented a database accessible through **AWS Athena**, with data cataloging facilitated by **AWS Glue's crawler**.

[Customer Churn predictive Analysis](#)

*Jupyter Notebook, Pandas, Numpy*

- Focused on predicting customer churn for a bank, addressing an imbalanced dataset with the **SMOTE technique**.
- Achieved an impressive 10% improvement in overall accuracy and significant enhancements in precision and **F1\_score** using cross validation set. Explored various machine learning models, including **logistic regression, K-nearest neighbors (KNN), (SVC), decision trees, and random forests**.

[MongoDB and NodeJS - Blog app](#)

*Node JS, Mongo DB, Express, Vanilla JS, CSS*

- Developed a dynamic server-side rendering application using Node.js.
- Utilized EJS templates and MongoDB schemas for efficient data handling.

## Academics

MS Computer science - Harrisburg University of Science and Technology

Present

Bachelor of Science Computer Science

January 2021 - May 2023

Purdue University Fort Wayne - Senior Capstone Competition Project Winner 2023

# Maneth Balasooriya

Phone: 260-580-6204 | <https://github.com/manethportfolio>

Email: manethbalasooriya11@gmail.com

Objective: A results-driven and analytical software engineer and data enthusiast with a versatile skillset and a passion for the software industry. I'm seeking a challenging position to utilize my technical expertise, problem-solving abilities, and collaborative mindset to contribute to the success of a dynamic software engineering team.

## Education:

**Masters in Computer science** May 2025  
Harrisburg University of science and Technology

**Bachelor of Science in Computer Science** May 2023  
Purdue University Fort Wayne Indiana  
Senior Capstone Project Winner 2023

## Skills:

- Statistical Analysis: Proficiency in statistical methods for drawing insights from data, including hypothesis testing, regression analysis, and ANOVA.
- Programming: Strong programming skills in languages like Python or R, which are commonly used for data analysis and modeling.
- Data Cleaning and Preprocessing: Skill in cleaning and preparing data for analysis, addressing missing values, outliers, and inconsistencies.
- Machine Learning: Understanding of machine learning algorithms and the ability to build, train, and evaluate predictive models for various tasks.
- Database Management: Knowledge of database systems, including SQL for data retrieval, manipulation, and management.
- Data Visualization: The ability to create compelling and informative data visualizations using tools like Matplotlib, Seaborn, Tableau, or Power BI.
- Business Acumen: A solid understanding of the industry or domain in which you work, enabling you to relate data insights to business objectives.
- Communication Skills: Strong written and verbal communication skills to effectively convey findings and insights to both technical and non-technical stakeholders.
- Problem-Solving: The ability to approach complex problems logically and find innovative data-driven solutions.

## Projects:

- MapAuthority: Developed a web application using Javascript and MongoDB to streamline data management processes, bringing back a

seller platform back to life which is hosted and managed on AWS using the SP-API.

- Movie recommendation system: Implemented automated testing procedures using React JS and machine learning algorithms, resulting in a **30%** reduction in software bug and a significant change in the functionality of the application.

**References/Referrals** : Available upon request

**Personal Projects:**

- Data Pipeline: Created a robust real-time data pipeline using Kafka and AWS S3 for efficient data ingestion and storage, and established a user-friendly database accessible through AWS Athena. Simplified data cataloging through AWS Glue's crawler, streamlining data management.
- ML (Machine Learning): Managed a project focused on forecasting customer churn in the banking sector. Employed the SMOTE technique to address dataset imbalances, resulting in an impressive **12%** accuracy improvement and notable enhancements in precision and F1 scores. Evaluated multiple machine learning models, including logistic regression, K-nearest neighbors, support vector machines, decision trees, and random forests, ensuring a comprehensive analysis for model selection

I am excited about the opportunity to contribute to a software engineering team and apply my skills to deliver high-quality software solutions. Please feel free to contact me for further discussion or to request references.

Sincerely,  
Maneth Balasooriya

# Christine Park

Fairfax, VA | +1 571-439-2333 | [cspark58@jhu.edu](mailto:cspark58@jhu.edu) | [www.linkedin.com/in/cspark58](https://www.linkedin.com/in/cspark58)

## EDUCATION

---

### Johns Hopkins University

Bachelor of Science: Computer Science, Minor in Applied Mathematics and Statistics

Baltimore, MD

Expected May 2025

Cumulative GPA: 3.81/4.0; Dean's List 2021-2022

Relevant Coursework: Gateway Computing: Python, Intermediate Programming, Data Structures, Mathematical Foundations for CS, Full-Stack Javascript, Intro to Algorithms, Intro to Data Science, Computational Genomics: Sequences

## SKILLS

---

**Languages:** Korean, Mandarin

**Technical:** Python, R, Java, Javascript, C, C++, SQL

## WORK EXPERIENCE

---

### Johns Hopkins University

Sept 2023 – Present

Elbert's Lab, Data Science Intern

- Integrated PIDINST data schema to enable automated linking and the assignment of unique identifiers to instruments for consistent and accurate data management
- Implemented cross-data integration to foster broader data accessibility and collaborative use across platforms
- Partnered with researchers at Hopkins Extreme Materials Institute (HEMI) and the Institute for Data Intensive Engineering & Science (IDIES) to enhance interdisciplinary research initiatives, optimize data acquisition processes, and bolster data management practices

### Johns Hopkins School of Medicine

Mar 2023 – Present

Larman Lab, Analytics Research Intern

- Streamlined the project's analytical phase by implementing a robust R-based pipeline for genomic sequence interpretation, resulting in a 20% time reduction in the process
- Constructed data visualization techniques to effectively communicate complex evolutionary relationships
- Developed a comprehensive enterovirus genomic database, functioning as an invaluable resource for the research team

### Johns Hopkins Applied Physics Laboratory

Jun 2023 – Aug 2023

Software Engineer Intern

- Employed Matplotlib, NumPy, Trimesh, and SciPy to develop dynamic and interactive visualization tools for the comprehensive analysis of asteroid rubble piles for NASA's Double Asteroid Redirection Test (DART) mission
- Integrated multi-source data sets to generate accurate 3D models of asteroid structures, enhancing mission planning and risk assessment
- Communicated with engineers and scientists to align objectives and deliver actionable insights that informed mission-critical decisions

## COMMUNITY

---

### Gateway to Python Course Assistant

Aug 2023 – Present

- Guided 50+ students by answering in-class questions, grading assignments, and providing one-on-one tutoring while collaborating with professors on biweekly quizzes

### HULA (Hopkins Undergraduate Leaders in Advising) Peer Mentor

July 2022 – Present

- Mentored 20+ first-generation and limited-income first-year students at Johns Hopkins University, facilitating their smooth transition through biweekly meetings to address individual concerns

### Kappa Kappa Gamma Academic Excellence Director

Jan 2022 – Present

- Collaborated with executive board to elevate academic standards and supported members' success through recognition, open study hours, and care packages during exams to cultivate a close-knit community

# Trisha Karani

Baltimore MD | tkarani1@jhu.edu | (510) 371-1895 | [linkedin.com/in/trishakarani/](https://linkedin.com/in/trishakarani/) | [github.com/tkarani1](https://github.com/tkarani1)

## EDUCATION

### Johns Hopkins University

Bachelors of Science in Computer Science and Biomedical Engineering | **GPA 3.68/4.0**

Baltimore, MD

May 2023

M.S.E in Computer Science

Expected May 2025

**Relevant Coursework:** Machine Learning, Deep Learning, Artificial Intelligence, Databases, Object-Oriented Software Engineering, Algorithms, Data Structures, Data Science, Computational Genomics, Mobile App Development, Probability and Statistics

## SKILLS

**Technical Languages:** Java, Javascript, Python, C, C++, Swift, MATLAB, HTML, CSS

**Frameworks:** sleep.ai, Python Notebook (PyTorch, NumPy, Matplotlib, Scikit-learn), Full-stack (MongoDB, Node.js), React, MariaDB, MySQL, SQL, Linux/Unix, Git

## EXPERIENCE

### FiOR Medical | Biomedical Engineering Department Design Team – *Design Team Co-founder*

March 2021 - Present

- Prototyped an automated sensor-based fire-prevention device in **Arduino** to reduce the incidence of operating room fires
- Designed and conducted research experiments to plot the relationship of oxygen sensor readings and surgical oxygen levels
- Conducted 10+ stakeholder interviews and market viability research to determine product demand and feasibility
- Filed a provisional patent and conducted design verification testing at Johns Hopkins Hospital

### Knee Implantable Device | Biomedical Engineering Department Design Team – *Project Lead*

March 2022 - June 2023

- Designed, and developed a prototype for an implantable sensor-based strain measuring device to provide doctors with metrics for knee recovery monitoring
- Established an agile development framework for an 8-person undergrad team by creating milestones, and running weekly meetings
- Explored medical and market research by fostering clinical collaborations

### Mayo Clinic | Applied Computational Neurophysiology and Neuromodulation Laboratory – *Summer Researcher*

June 2021 – August 2021

- Created a pipeline for developing a **deep learning** by training sleep.ai ML models using mp4 video clip files of rat treadmill movements, and using labeled frames to track limb movements across frames
- Developed **Shell scripts** based automation to run training, and learning for prediction of 100+ videos
- Developed **Python notebooks** with numpy, matplotlib for extracting, calculating, and creating visualization plotting KPIs (position data, swing length, stance duration) to analyze rat gait for an induced Parkinsonian lesion

### Institute for Computational Medicine | Computational Cardiology Lab – *Undergrad Researcher*

March 2020 – May 2021

- Analyzed patient-specific 3D virtual heart models using OpenCarp to simulate action potentials and determine optimal ablation sites
- Developed computational models of 2D heart tissue models to simulate heart attacks with varied ionic properties using **MATLAB** to understand the impact of repolarization gradients on ventricular tachycardia morphology

## PROJECTS

### Music Database – *Project Member*

January 2022 – May 2022

- Used entity-relationship data modeling to structure a database of songs and geo-demographics on **relational SQL (MariaDB)**
- Created frontend platform that allowed users to answer various song-related questions that were answered by SQL queries using **PHP** and **HTML** and visualized graphically with **CanvasJS**
- Optimized queries using prepared statements and developed visualizations (including tables and graphical reports)

### Convolutional Neural Network for ASL Classification – *Project Member*

January 2022 – May 2022

- Created a **deep learning model** that takes images of ASL hand symbols and provides classification for letters of the English alphabet
- Implemented a **convolutional neural network (using PyTorch, Scikit-learn)** that includes preprocessing, data augmentation (Gaussian blur, translate, rotate, etc) and feature engineering (convexity, etc) to create a prediction with 99% accuracy

### HackNotes Virtual Runnable Container – *Project Member*

August 2021 – December 2021

- Designed and developed an online note-taking **Full-stack web application portal (using Node.js, React, Javascript, MongoDB)** with managed user sessions
- Implemented virtual machine containers to run code snippets and utilizing ML algorithm for image-to-code conversion

## LEADERSHIP

### JHU Computer Science Department | Intermediate Programming (C/C++), Data Structures – *Teaching Assistant*

August 2020 - Present

- Provided in-class support for programming exercises and understanding recursion, memory allocation, algorithm design, etc

# PRADEEP BOLLEDDU

+1 774-635-8035 | [bpradeep@umassd.edu](mailto:bpradeep@umassd.edu) | [Website](#) | [Github](#) | [Linkedin](#)

Experienced Business Intelligence Analytics developer proficient in statistical Data Analytics, Data Models, SQL, and data visualization tools, showcasing strong communication skills, Microsoft Excel, and PowerPoint expertise. Currently Pursuing a Master's degree in Data Science. **Sponsorship Not Required, Eligible to Work in USA.** [Merck: Internship 2024.]

## EDUCATION

---

### University of Massachusetts – Dartmouth

Master's, *Data Science with Concentration in Artificial Intelligence*

Course Work: *Data Engineering, Advance Statistics, Probability, High performance computing & Cloud computing.*

January 2023 - December 2024

GPA: 3.8

### Indian Institute of Technology Ropar (IIT)

Bachelor's, *computer engineering*

August 2019 - March 2023

GPA: 3.5

## INTERNSHIP EXPERIENCE

---

### Semusi Technologies Private Limited – Sales Force Client

Delhi, India

Big Data Product Analyst

April 2021 - July 2022

1. Designed and created **10+** dashboards for cross functional stakeholders to track data annotation progress.
2. Developed complex and reusable SQL queries to quantify annotation quality metrics across **1000+** annotation job queues & Lead Power BI dashboard within the business unit to monitor key performance.
3. Created robust business intelligence solutions for the operations teams, Scrape data from Azure APIs to feed modeling and dashboards. Identify anomalies through automated scripts and Performed A/B Testing with Ad hoc Analysis.
4. Tested and verified the behavior of ML systems And Retrieved & updated documents in MongoDB.

### Techno Colab's software solutions Inc

Mumbai, India

Machine Learning Engineer

September 2020 - February 2021

1. Improved existing workflows and built a data system from scratch, capable of handling large scale variable data using Spark and Cassandra bringing down system latency to under **10** second.
2. Developed and trained a CNN for a startup company to classify over **2000** different medical symptoms with more than **87%** of accuracy.
3. Performed clustering and derived insights about customer behavior from their usage patterns on AWS & Designed an interface for Apache Spark to enable customers to access datastores on multiple platforms.
4. Designed, developed and deployed recommender systems for company platforms to drive business results.

## SKILLS & INTERESTS

---

- **Programming Languages:** Python, Java, C++, Go/Golang, MySQL & R-Programming.
- **Frameworks:** Agile, NLP, CI/CD, Tensor flow, Keras, Py torch, LINUX, Scrum &OpenCV.
- **Cloud Services:** Azure, ETL, Apache Spark, airflow, cloud infrastructure, Tableau, Power BI, VBA &MongoDB.
- **Technologies:** Git, AWS, SAP, Snowflake, Hive, Spark, HDFS, VBA, Kafka &data visualization tools.
- **Data base:** Postgres SQL, lambda, MongoDB, S3, Informatica, SQL Server & Data analytics.
- **Tech Stack:** Algorithms, A/B testing, Communication skills, Cutting, Data-driven, Entrepreneurial, Integrity, Lean & Teamwork.

## PROJECTS

---

### 1. Real-time Log Query Interface for large data sets using Apache Spark | [Github link](#)

- Log Query Interface is a web application enabling efficient dataset querying through user input on the website, with significant response time enhancements via server-side Spark clusters for distributed dataset storage.

## RESEARCH PUBLICATIONS

---

### 1. Analysis of Clustering Techniques | [Paper Link](#) | ISSN: 2456-3307 / IJSRCSEIT / Volume 8, Issue 2

- Research investigated various clustering techniques, such as K-means, DBSCAN, Agglomerative Hierarchical, and Gaussian Mixture models, to analyze high-dimensional credit card customer data for segmentation and marketing strategies, emphasizing the intricacies of handling such data.

### 2. Hadoop: An Overview of Data security | [Paper link](#) | ISSN: 2319-7064 / IJSR / Volume 11 Issue 3, March 2022

- Experiment involved setting up Ubuntu Linux 12.04, configuring various components such as Hadoop, OAuth authentication, and encryption methods. The results indicated that encryption using OAuth memory performed as a more efficient and secure encryption method, highlighting the importance of encryption in Hadoop clusters for data processing

## POSITION OF RESPONSIBILITIES

---

### 1. Head Coordinator – Finance & Economics Club

### 2. Mentor – The Artificial Intelligence Club

# **Qijia HE**

Email: [heqj3@uw.edu](mailto:heqj3@uw.edu) | Phone: +1 (206) 503-7975

## **RESEARCH INTERESTS**

Causal inference; Data-driven decision making; Machine learning

## **EDUCATION**

<b>University of Washington</b>	Seattle, WA
Ph.D. in Statistics	09.2023 - Present
M.S. in Statistics, GPA: 3.90/4.0	09.2021 - 03.2023
<b>Sun Yat-sen University</b>	Guangzhou, China
B.S. in Statistics, GPA: 3.91/4.0	08.2017 - 06.2021
Scholarships: First-Class Scholarships, SYSU (2020, 2019), Second-Class Scholarship, SYSU (2018)	
<b>University of California, Berkeley</b>	Berkeley, CA
Exchange student, GPA: 4.0/4.0	08.2019 - 12.2019

## **PUBLICATIONS**

- He, Q.**, Gao, F., Dukes, O., Delany-Moretlwe, S., Zhang, B. (2023). Generalizing the intention-to-treat effect of an active control from historical placebo-controlled trials to an active-controlled trial. arXiv preprint arXiv:2304.03476. Under review at *Journal of the American Statistical Association*.
- He, Q.**, Zhang S., LeBlanc, M. L., Zhao, Y. Q. "Estimating individualized treatment rules by optimizing the adjusted probability of a longer survival." Under review at *Statistical Methods in Medical Research*.

- Book Chapter:

- He, Q.**, Zhao, Y. Q. (2023). "Statistical Learning Methods for Estimating Optimal Individualized Treatment Rules from Observational Data." In: Chakraborty, B., Moodie, E.E., Laber, E.B., Cai, T., van der Laan, M, editors. *Handbook of Precision Medicine (Under review)*
- Research on the Development Trend and Social Effect of Digital Economy* (In Chinese), published by China Social Sciences Press (ISBN: 978-7-5203-7735-5), engaged in the writing of all chapters except Chapter 4

## **RESEARCH EXPERIENCE**

<b>Fred Hutchinson Cancer Research Center</b>	Seattle, WA
Generalizing the intention-to-treat effect (ITT) from historical trials to a planned non-inferiority (NI) trial	08.2022 - 04.2023
Advised by Prof. Bo Zhang, Prof. Fei Gao	
<ul style="list-style-type: none"><li>Collaborated in the novel design of a causal inference framework to infer the ITT effect of the active control using relevant historical placebo-controlled trial data</li><li>Designed and implemented simulation studies to evaluate the proposed estimators under different identification assumptions</li><li>Obtained historical-data-driven estimates under point/partial identification and discuss various strategies for sensitivity analyses in the application</li></ul>	

Validating individualized prevention rules (IPRs) using post-randomization events for HIV prevention 08.2022 - 09.2023

Advised by Prof. Ting Ye (University of Washington), Prof. Fei Gao, Prof. Bo Zhang

- Developed an IPR from a classification perspective using survival data from HPTN (HIV Prevention Trials Network) with the goal of optimizing the population-level HIV acquisition outcome in a limited resource setting
- Develop a CAB-LA recommendation score (CAB-RS) to assist physicians' decision-making in prescribing CAB-LA or daily oral TDF/FTC based on a person's risk factors in real time

Optimal adjusted probability learning for individualized treatment rules with censored data

03.2022 - 11.2022

Advised by Prof. Yingqi Zhao

- Proposed a new criterion of treatment effect for constructing ITRs that optimize clinical benefits with survival outcomes
- Developed a new optimal ITR estimation method by maximizing a nonparametric estimator of the proposed criterion
- Applied the proposed method to four data simulations and performed data analysis using data collected from SWOG S0819, outperforming other recently developed methods in terms of accuracy and (or) value function

**Sun Yat-sen University**

Guangzhou, China

Approximate covariate balancing for causal inference

08.2020 - 10.2020

Advised by Prof. Ying Yan

- Implemented the proposed weighting method, which approximately balances covariates by optimizing an entropy loss under an inequality constraint constructed by a generalized Mahalanobis distance, through coordinate descent via R
- Proved that the dual problem of the proposed weighting method is an  $l_2$  regularized unconstrained optimization problem by applying Fenchel duality theorem

Semi-supervised pattern recognition algorithm with label noise based on SVM-KNN

03.2020 - 08.2020

Advised by Prof. Jia Li

- Reformed KNN to build the regularization model with weighted quadratic loss function and gradient descent
- Implemented the proposed four-stage semi-supervised algorithm based on KNN and SVM that includes denoising, initialization, updating, and cross prediction
- Applied the algorithm to AR-Face database via Matlab, reaching accuracy of above 90% with only 1% known labels and 10% random label noise

## PRESENTATIONS

---

Generalizing the intention-to-treat effect of an active control from historical placebo-controlled trials to an active-controlled trial

- The Translational Data Science Integrated Research Center Retreat, Poster Presentation. Kirkland, WA, 2023.
- 20th Annual STI & HIV Research Symposium, Poster Presentation. Seattle, WA, 2023.

## TEACHING AND TUTORING EXPERIENCE

---

**Academic tutoring center, School of Mathematics, Sun Yat-sen University**

09.2018 - 12.2018

Tutor in Mathematics

- Assisted two international students from School of Business with Calculus for two hours per week

**TAL Education Group**

01.2018 - 05.2019

TA in primary-school Olympiad Mathematics

- Prepared coursework solutions, graded homework assignments, and held TA sessions

## SKILLS

---

R, Python, MATLAB, C++, Latex



# JADEN REEVES

## CLOUD COMPUTING

[✉ reevesjaden3@gmail.com](mailto:reevesjaden3@gmail.com) [\(443\)-301-2588](tel:(443)-301-2588)

[in https://www.linkedin.com/in/jaden-reeves-35b836249/](https://www.linkedin.com/in/jaden-reeves-35b836249/) [G https://github.com/](https://github.com/)

## EDUCATION

### Bachelor of Science in Cloud Computing

Morgan State University, Fall 2022 - Spring 2026 (Expected)

Fall 2022 - Spring 2026

- Intro to Computer Science 1 & 2

- Data Structures & Algorithms

- Computer Systems & Digital Logic

### Western School of Technology - IT Networking

Fall 2018 - Summer 2022

## TECHNICAL SKILLS

**CODING LANGUAGES:** Python

**MICROSOFT SUITE APPLICATIONS:** Microsoft Word, Microsoft PowerPoint, Microsoft Excel

## PROJECTS

### FINANCIAL LITERACY AND WELLNESS APPLICATION: CRED - ED

Apr. 2023 - Apr. 2023

Created a prototype application during the CodeLinc Hackathon 2023 - Hosted by Lincoln Financial Group to deliver specific information to the user about financial literacy and financial wellness pertaining to their current situation in life, with a gamified aspect of badges earned through accolades like consistency.

### HANG-MAN GAME: ANIMALS

Dec. 2022 - Jan. 2023

A Hang-Man game created using if statements until a while loop to continue running the game. Your lives are counted as the game progresses, while keeping track of previously guessed letters.

### LEGENDARY ADVENTURER BATTLE

Dec. 2022 - Jan. 2023

Created a randomized battle simulator with the users choice of name and race (human, elf, etc.) Utilizing import of previously created modules and newly created functions, I created a roll dice function that randomly chooses a number on a specified number of sides. Using python specific keywords, I created a while loop to continuously run my code within the chosen number of rounds.

### ADVENTURE GAME

Oct. 2022 - Nov. 2022

An adventure game to save the princess with multiple routes of success. This is done by utilizing if else statements, nesting multiple if statements, and elif statements.

### MARVEL TRIVIA GAME

Sept. 2022 - Current

A trivia game that tests the users knowledge of marvel comics using if else statements, while the code prints out 4 answer choices pick from.

### 100 DAYS OF CODE CHALLENGE- REPLIT

Sept. 2022 - Current

Daily replit coding challenges to learn new coding skills on a steady pace.

## ACTIVITIES

### **Morgan State University 2022 CodeLinc Hackathon, Competitor**

Apr. 2022 - Apr. 2022

Lincoln Financial Hosted Hackathon: Created the Cred - Ed app in my projects.

### **Certificate of Membership for NAACP, General Body Member**

Oct. 2022 - Dec. 2022

NAACP member at Morgan State University.

## EMPLOYMENT

**Giant Foods**, 10210 Mill Run Circle Owings Mills, MD  
*Starbucks barista*

May 2022 - Aug. 2023

Innovative and results-driven Computer Science professional with a strong focus on AI and Data Science. Proven track record of leveraging deep technical expertise to develop intelligent solutions that enhance business efficiency and drive innovation. Recognized for leadership in academic research and commitment to ethical data practices.

## EDUCATION

---

- **San Francisco State University (SFSU)** San Francisco, CA  
*Master of Science in Computer Science with an emphasis on AI; CGPA: 3.75*  
*Anticipated: August 2024*
- **Gujarat Technological University** Ahmedabad, India  
*Bachelor of Engineering in Computer Engineering; CGPA: 3.7*  
*August 2020*

COURSEWORK: Data Structures, Analysis of Algorithms, AI Ethics and Explainability, Data Visualization, Big Data Analytics, Probabilistic ML, Reinforcement Learning, Database Management Systems, Multivariate Calculus

CERTIFICATIONS:

- **Data Science Certificate** — DataCamp
- **Deep Learning Specialization** — DeepLearning.AI
- **AI Policy Certification** — Center for AI and Digital Policy

## SKILLS

---

- **Languages:** Python, SQL, Java, C++, MATLAB, R, Javascript, Scala, Regex
- **Technologies:** Git, AWS, Azure Cloud, React, Apache Spark, Hadoop, Jupyter Notebook, Tableau, PowerBI, LaTeX
- **Libraries:** Pandas, Numpy, Scikit-learn, TensorFlow, Pytorch, Matplotlib, Seaborn, Plotly, PySpark, SciPy, SpaCy
- **Data Science:** Data Management and Governance, Exploratory Analysis, Model Fitting, Data Visualization, Storytelling

## EXPERIENCE

---

- **Language-Based AI Lab at SFSU** San Francisco, CA  
*Research Engineer*  
  - Boosted logical reasoning accuracy of LLMs by creating specialized datasets and iterative prompt engineering.
  - Orchestrated development of intelligent agents, by fine-tuning open-source foundation models, that integrate tools to enhance critical thinking capabilities. **Tools used:** Pytorch, HuggingFace Transformers, QLORA, LangChain, wandb.
  - Mentored undergraduates working on NSF-funded research to benchmark capabilities of LLMs on a spectrum of tasks.
- **Associated Students at SFSU** San Francisco, CA  
*Member, Board of Directors*  
  - Mastered parliamentary procedures to streamline institutional decision-making, enhancing governance efficiency.
  - Amplified student engagement twofold by strategizing and launching a university-wide opportunities board.
  - Fostered interdisciplinary collaboration through continuous partnerships with various university departments.
- **VTech Elite** Gujarat, India  
*Software Engineer*  
  - Developed a scalable resource management platform. **Tools used:** Python, React, Angular, Node, Express, MySQL.
  - Elevated user satisfaction by 70% through user-centric A/B testing and UI redesign of analytics software.
  - Catalyzed a 55% revenue increase by integrating a personalized recommender system which improved client retention.
- **Parality XR** Gujarat, India  
*Software Engineer Intern*  
  - Developed AR and VR based software for an educational application. **Tools used:** Unity Engine, Blender and C#.
  - Enhanced deployment speed by 70%, ensuring on-time delivery through proficient use of Git and Azure pipelines.
  - Accelerated learning engagement by 150% after conducting controlled trials with students to optimize the application.

## PROJECTS

---

- **AI Ethics Case Study** Fall 2022
  - Led a critical examination of an existing AI application using model explainability and through algorithmic audits.
  - Produced a comprehensive report detailing model risks, to showcase ethical AI considerations within the industry.
- **Quantum Machine Learning** Fall 2022
  - Implemented a Quantum Kernel-based SVM classifier, demonstrating its viability on quantum computers.
  - Spearheaded the development of a Quantum Computing course for ML that uses interactive Python Notebooks.

# Daniel Herbert

14 Blue Rock Drive Sellersville, PA 18960 | 267-567-5175 | djh6502@psu.edu

## Objective

Motivated and ambitious college sophomore, majoring in Energy Business and Finance (EBF), seeking internships or entry-level positions to gain practical experience and contribute to a company's priorities.

Committed to utilizing my knowledge in energy markets, financial analysis, and sustainable business practices to drive growth and profitability. Dedicated to further developing skills in project management, data analysis, and strategic decision-making while expanding my professional network.

## Education

**Penn State University**

Major: Bachelor of Science, Energy Business Finance

May 2026

## Relevant Coursework

- Meteo 3 - Meteorology
- Geosc 40 - The Sea Around Us
- CMPSC 101 - Intro to Programming
- Econ 102 – Microeconomics
- Econ 104 – Macroeconomics
- CAS 100A - Effective Speech
- Math 22 - Algebra 2
- Math 26 - Trigonometry
- Math 140 - Calculus 1
- Math 141 - Calculus 2

## EXPERIENCE

Coordinator, delivery and take-out

May 2022 - Present

**Perkasie Pizza & Pasta Restaurant** | Perkasie, PA

- Managed both walk-up and phone orders, ensuring smooth and efficient order management and customer service
- Developed effective communication skills by coordinating with delivery drivers to ensure on-time pickups and deliveries for a seamless customer experience.

Team Member

September 2021 - May 2022

**Saint Rocco's Treats** | Perkasie, PA

- Assisted a local dog food company in managing inventory and stock control, gaining hands-on experience in the production process.
- Participated in local community events to market the company's business, engaging with potential customers, and showcasing products and services.

Pool Manager

May 2021 – August 2021

**Orchard Hill** | Hilltown, PA

- Acted as a point of contact for patrons, addressing any complaints or needs promptly and professionally, enhancing overall customer satisfaction.
- Assisted in the direction and supervision of lifeguards, ensuring they were positioned effectively and following established safety procedures.

Cook

January 2020 - May 2021

**Red Robin Restaurant** | Quakertown, PA

- Demonstrated expertise in food handling and storage procedures to ensure proper food safety and quality.
- Collaborated with the kitchen and serving team to ensure quality of customer service.

## VOLUNTEER

Brookdale Memory Care

January – December 2020

# LUAN LAM

Houston, TX • (713) 367-5859 • luanlam97@gmail.com

## EDUCATION

### GEORGIA INSTITUTE OF TECHNOLOGY – Atlanta, GA

#### *Master of Science in Computer Science*

Expected December 2024

- Specializing in Machine Learning

### UNIVERSITY OF TEXAS AT AUSTIN – Austin, TX

#### *Bachelor of Arts in Economics*

Graduated January 2020

- Business Foundation certificate
- Applied Statistical Modeling certificate

- Risk Management certificate
- Elements of Computing certificate

## EXPERIENCE

### DISH WIRELESS – Littleton, CO

May 2023 - August 2023

#### *Data Science Intern*

- Developed a modular python package that utilities ffmpeg and opencv to optimize pre/postprocessing videos and leverages pytorch to facilitate deep learning functionality
- Deployed Yolov8, a real-time object detection in production environments, utilizing tools like GitHub workflow, Docker, and cloud platforms such as AWS Lambda to optimize models for performance and scalability
- Utilized cloud-based machine learning platform, AWS Sagemaker, to streamline the model training and deployment workflow, enabling faster iteration and experimentation
- Published package on PyPI platform to enable easy installation and distribution to a wide range of users and environments
- Conducted comprehensive research on the theory and practical application of diffusion models to lead the team on the develop a customized model tailored to specific use-cases

## PROJECT

### Multi Agent Soccer Learner – Reinforcement Learning

January 2023 – May 2023

- Utilized various RL algorithms such as A3C, DQN and PPO to facilitate the training of multiple agents in a simulated soccer game
- Leveraged Python and popular libraries like OpenAI Gym and Ray to design and train RL agents, achieving significant performance improvements and robustness

### Stock Price Predictor – Machine Learning for Trade Project

September 2022 – December 2022

- Created and implemented a dynamic market simulator to rigorously evaluate the effectiveness of popular trading indicators and techniques
- Utilized Sci-kit learn and Pandas library to design and deploy a diverse range of learners, including decision trees, bagging, q-learner, and neural network to capture stock price movements

### Job Offer Management – Android Application Project

September 2022 – December 2022

- Created a responsive android app where users can upload and compare their job offers against other users
- Managed database using SQLITE, developed back-end with Java and used Android Studio for UX design

## SKILLS

**Programming Languages:** Python, C++, SQL, JavaScript, HTML/CSS

**Developer Tools:** Git, Docker, Hugging Face, AWS Lambda, AWS Sagemaker, GCP

**Libraries:** OpenCV, Pandas, NumPy, Pytorch, SkLearn, D3, Ray-rllib,

**Certificates:** AWS CCP, AWS Develop Associate (in progress)

**Other:** Agile development, Object-Oriented Programming, Jira

# Tom Earnest

Email: [tom.earnest@wustl.edu](mailto:tom.earnest@wustl.edu)

Phone: 262-902-5020

Website: <https://earnestt1234.github.io/>

Github: <https://github.com/earnestt1234/>

LinkedIn: <https://www.linkedin.com/in/earnestt1234/>

## EDUCATION

---

*Doctor of Philosophy, Computational and Data Sciences* 2025 (expected)  
Division of Computational and Data Sciences  
Washington University in St Louis

*Master of Science, Psychiatric Research* 2017  
Institute of Psychiatry, Psychology, and Neuroscience  
King's College London

*Bachelor of Arts, Biology with Concentration in Neuroscience* 2016  
Grinnell College

## RESEARCH EXPERIENCE

---

### Doctoral Research 2020-current

Washington University in St. Louis

Advised by [Aristeidis Sotiras, Ph.D.](#)

- Used data-driven methods (especially non-negative matrix factorization) to define and evaluate a staging system for brain pathology in Alzheimer's Disease
- Applied supervised machine learning with nested cross-validation for prediction of cognitive decline in Alzheimer's Disease
- Managed storage and organization of three multimodal neuroimaging datasets in a high-performance computing cluster
- Developed code for preprocessing, analysis, and visualization of neuroimaging data from >1,000 subjects

### Research Technician II 2019-2020

Washington University in St. Louis

Supervised by [Alexxai Kravitz, Ph.D.](#), and [Meaghan Creed, Ph.D.](#)

- Created two graphical user interfaces in Python for analysis and visualization of data from open-source rodent experimentation devices
- Developed code for visualization of rodent electrophysiological data
- Collected behavioral and neuroanatomical data to study reward and motivation circuits in rodents

## **Postbac Program**

2017-2019

National Institute of Dental and Craniofacial Research

Supervised by Hans Jürgen Solinski, Ph.D., and Mark Hoon, Ph.D.

- Learned programming to automate research tasks, such as experimental blinding and cell counting
- Ran behavioral experiments to study the neurobiology of sensorimotor circuits in mice
- Bred and genotyped transgenic mice
- Used confocal imaging to characterize neuronal populations involved in pain and itch

## **Master's Research**

2016-2017

Institute of Psychiatry, Psychology, and Neuroscience; King's College London

Supervised by Elizabeth Shephard, Ph.D., and Patrick Bolton, Ph.D.

- Primary researcher on a project using movement trackers to assess hyperactivity in a population of children and adolescents with tuberous sclerosis complex
- Traveled within England to administer psychometric assessments to individuals with varying levels of cognitive and behavioral impairment (~30 visits)
- Developed statistical analyses of motion tracking data to predict cognitive constructs

## **Mentored Advanced Project**

2015-2016

Grinnell College

Supervised by Nancy Rempel-Clower, Ph.D.

- Co-planned and ran a project studying the effects of stress on behavior and neurobiology in adolescent rodents
- Applied tests to study anxiety-like behavior in rats
- Conducted brain dissections to measure neuroanatomical changes correlated with behavioral differences

## **PUBLICATIONS**

---

*Leading asterisk signifies preprint or submitted work.*

Bani, A., Ha, S. M., Xiao, P., **Earnest, T.**, Lee, J., & Sotiras, A. (2023). Scalable Orthonormal Projective NMF via Diversified Stochastic Optimization. In A. Frangi, M. de Bruijne, D. Wassermann, & N. Navab (Eds.), *Information Processing in Medical Imaging* (pp. 497–508). Springer Nature Switzerland. [https://doi.org/10.1007/978-3-031-34048-2\\_38](https://doi.org/10.1007/978-3-031-34048-2_38)

De Araujo Salgado, I., Li, C., Burnett, C. J., Rodriguez Gonzalez, S., Becker, J. J., Horvath, A., **Earnest, T.**, Kravitz, A. V., & Krashes, M. J. (2023). Toggling between food-seeking and self-preservation behaviors via hypothalamic response networks. *Neuron*, 111(18), 2899-2917.e6. <https://doi.org/10.1016/j.neuron.2023.06.006>

\* **Earnest, T.**, Bani, A., Ha, S. M., Hobbs, D. A., Kothapalli, D., Yang, B., Benzinger, T. L. S., Gordon, B. A., & Sotiras, A. (submitted). Data-driven decomposition and staging of flortaucipir uptake in Alzheimer's Disease. *Alzheimer's & Dementia*.

**Earnest, T.**, Shephard, E., Tye, C., McEwen, F., Woodhouse, E., Liang, H., Sheerin, F., & Bolton, P. F. (2020). Actigraph-Measured Movement Correlates of Attention-Deficit/Hyperactivity Disorder (ADHD) Symptoms in Young People with Tuberous Sclerosis Complex (TSC) with and without

Intellectual Disability and Autism Spectrum Disorder (ASD). *Brain Sciences*, 10(8), Article 8. <https://doi.org/10.3390/brainsci10080491>

\* Kumar, S., **Earnest, T.**, Payne, P. R. O., Sotiras, A., & Initiative, the A. D. N. (2023). *Analyse patient-level heterogeneity in Alzheimer's Disease using multimodal normative modelling* (p. 2023.08.15.553412). bioRxiv. <https://doi.org/10.1101/2023.08.15.553412>

Matikainen-Ankney, B. A., **Earnest, T.**, Ali, M., Casey, E., Wang, J. G., Sutton, A. K., Legaria, A. A., Barclay, K. M., Murdaugh, L. B., Norris, M. R., Chang, Y.-H., Nguyen, K. P., Lin, E., Reichenbach, A., Clarke, R. E., Stark, R., Conway, S. M., Carvalho, F., Al-Hasani, R., ... Kravitz, A. V. (2021). An open-source device for measuring food intake and operant behavior in rodent home-cages. *eLife*, 10, e66173. <https://doi.org/10.7554/eLife.66173>

Shephard, E., McEwen, F. S., **Earnest, T.**, Friedrich, N., Mörtl, I., Liang, H., Woodhouse, E., Tye, C., & Bolton, P. F. (2022). Oscillatory neural network alterations in young people with tuberous sclerosis complex and associations with co-occurring symptoms of autism spectrum disorder and attention-deficit/hyperactivity disorder. *Cortex*, 146, 50–65. <https://doi.org/10.1016/j.cortex.2021.10.007>

Slivicki, R. A., **Earnest, T.**, Chang, Y.-H., Pareta, R., Casey, E., Li, J.-N., Tooley, J., Abiraman, K., Vachez, Y. M., Wolf, D. K., Sackey, J. T., Kumar Pitchai, D., Moore, T., Gereau IV, R. W., Copits, B. A., Kravitz, A. V., & Creed, M. C. (2023). Oral oxycodone self-administration leads to features of opioid misuse in male and female mice. *Addiction Biology*, 28(1), e13253. <https://doi.org/10.1111/adb.13253>

Solinski, H. J., Dranchak, P., Oliphant, E., Gu, X., **Earnest, T. W.**, Braisted, J., Inglese, J., & Hoon, M. A. (2019). Inhibition of natriuretic peptide receptor 1 reduces itch in mice. *Science Translational Medicine*, 11(500), eaav5464. <https://doi.org/10.1126/scitranslmed.aav5464>

Solinski, H. J., Kriegbaum, M. C., Tseng, P.-Y., **Earnest, T. W.**, Gu, X., Barik, A., Chesler, A. T., & Hoon, M. A. (2019). Nppb Neurons Are Sensors of Mast Cell-Induced Itch. *Cell Reports*, 26(13), 3561–3573.e4. <https://doi.org/10.1016/j.celrep.2019.02.089>

Vachez, Y. M., Tooley, J. R., Abiraman, K., Matikainen-Ankney, B., Casey, E., **Earnest, T.**, Ramos, L. M., Silberberg, H., Godynyuk, E., Uddin, O., Marconi, L., Le Pichon, C. E., & Creed, M. C. (2021). Ventral arkypallidal neurons inhibit accumbal firing to promote reward consumption. *Nature Neuroscience*, 24(3), Article 3. <https://doi.org/10.1038/s41593-020-00772-7>

## PRESENTATIONS

---

**Earnest, T.**, Bani, A., Ha, S. M., Kothapalli, D., Yang, B., Lee, J., Sotiras, A. (2023, July). *Detecting floratacupir signatures of Alzheimer's Disease with non-negative matrix factorization*. Poster at Organization for Human Brain Mapping 2023. Montreal, Canada.

**Earnest, T. W.**, Solinski, H. J., Kriegbaum, M. C., Tseng, P. Y., Gu, X., Barik, A., ... Hoon, M. A. (2019, May). *Nppb-neurons are sensors of mast cell-induced itch*. Poster at NIH Postbac Poster Day. Bethesda, MD.

**Earnest, T. W.**, Solinski, H. J., Kriegbaum, M. C., Tseng, P. Y., Gu, X., Barik, A., ... Hoon, M. A. (2019, April). *Nppb-neurons are sensors of mast cell-induced itch*. Poster at the NIDCR Fellows Retreat. Washington, DC.

**Earnest, T.** (2017, July). *Using actigraphy to measure ADHD symptoms in tuberous sclerosis complex*. Poster at the IoPPN Psychiatric Research MSc year end session. London, UK.

**Earnest, T.** (2016, February). *Behavioral and morphological effects of stress in adolescent rats*. Talk in the Biology Student Seminar Series at Grinnell College. Grinnell, IA.

**Earnest, T.**, Yetter, M. (2015, November). *Behavioral and morphological effects of stress in adolescent rats*. Talk in the Psychology Student Seminar Series at Grinnell College. Grinnell, IA.

Yetter, M., **Earnest, T.**, Rempel-Clower, N. (2015, October). *Acute corticosterone treatment increases anxiety and dendritic elongation and arborization in the orbitofrontal cortex in mid-adolescent but not early-adolescent rats*. Poster at the Faculty for Undergraduate Neuroscience at the Society for Neuroscience Annual Meeting. Chicago, IL.

**Earnest, T.** (2015, September). *Acute stress increases anxiety behaviors in mid-adolescent rats and may cause dendritic elongation & arborization in the orbitofrontal cortex*. Poster at Iowa State Neuroscience Research Day. Ames, IA.

## AWARDS

---

- Outstanding Poster Award at NIH Postbac Poster Day (2019)
- NIH Postbaccalaureate Intramural Research Training Award (2017)
- Dean's Medal (2017): Best overall performance in all postgraduate taught programs in the Institute of Psychology, Psychiatry, and Neuroscience
- Sir Robin Murray Prize (2017): Best overall performance in the Psychiatric Research MSc
- Honorable Mention for poster presented at the Iowa State Neuroscience Research Day (2015)
- Dean's List for all semesters at Grinnell College (2012-2016)
- Trustee Honor Scholarship at Grinnell College (2012-2016)

## TECHNICAL SKILLS

---

- Programming
  - Proficient: Python, R
  - Experienced: Bash
  - Familiar: MATLAB
- Software
  - Git/GitHub
  - SLURM
  - Graphpad Prism
  - SPSS
  - Microsoft Office

# Debanjan Goswami

Cell No. (352)440-3323, Email: [debanjangoswami.iem@gmail.com](mailto:debanjangoswami.iem@gmail.com), LinkedIn: <https://www.linkedin.com/in/debanjan-goswami> GitHub : Projects

<b>Academic Degree</b>	<ul style="list-style-type: none"><li>• Ph.D. in Computer Science (CGPA: 3.91/4.0) Florida State University(<a href="#">FSU</a>), Tallahassee, FL, USA</li><li>• Bachelor of Technology in Electronics and Communication Engineering West Bengal University of Technology(<a href="#">WBUT</a>), Kolkata, IND</li></ul>	[Jan, 2020 - present] Aug, 2016
<b>Research Interest</b>	<b>Computer Vision, Deep Learning, Data Mining , Databases</b>	
<b>Experience</b>	<ul style="list-style-type: none"><li>□ <b>Research Assistant</b>, FSU AI LAB, CS Department, <a href="#">FSU</a>, Tallahassee, FL.<ul style="list-style-type: none"><li>• Efficient Image search using Show and Tell model based on Encoder-Decoder neural network. Deep CNN is used as the encoder to visualize an image into a fixed-length representation and decode the representation using LSTM network into a Natural Language Description.</li><li>• Unsupervised Web Table Clustering and Classification: Large(~75M) Web-Table Dataset is clustered based on its source and category. Used PySpark in Apache Spark 2.3 environment to cluster and classify Web-Tables with an accuracy over 90 percent.</li><li>• Sentiment Analysis on Twitter Using R: Performed sentiment analysis by taking twitter feed with #brexit as the dataset. Picked up data from particular region and of particular timeframe. Pre-processed the data to remove twitter jargons and term matrix is built. Bing Lexicon is used and to map sentiments with the most frequently occurring words fetched by a word cloud.</li></ul></li><li>□ <b>Data Scientist (PhD Intern)</b>, Reliable Deployment Lab, <a href="#">HP Inc.</a> , United States [Oct, 2022 - Jan, 2023]<ul style="list-style-type: none"><li>- Worked part-time as a Data Scientist Intern in Reliable Deployment Lab on multiple Computer Vision and Machine Learning Projects.</li></ul></li><li>□ <b>PhD Intern</b>, Reliable Deployment Lab, <a href="#">HP Inc.</a> , United States [May, 2022 - Aug, 2022]<ul style="list-style-type: none"><li>• Worked as a part of the Reliable Deployment Lab in a Federated Machine Learning Project. The Primary responsibility of the project is to build a smart decision-making system which would decide on actions based on historical manual actions made by humans in the existing process across hundreds of servers. Concepts from Active Learning and Uncertainty estimation algorithms are used to select highly informative samples to generate a standard high optimized ML model which achieved an accuracy close to 90 percent.</li></ul></li><li>□ <b>Research Associate Intern</b>, Artificial Intelligence Lab, <a href="#">Hewlett Packard Labs</a>, United States [June, 2021 - Aug, 2021]<ul style="list-style-type: none"><li>• Worked on a project named as Deep Active Learning with <a href="#">ClimateNet</a> Image Segmentation dataset where various uncertainty quantification(UQ) baselines such as Probabilistic ensembles, Batch Ensembles, Deep Ensembles, HyperBatch Ensembles and HyperDeep Ensembles were used to determine the uncertainty scores for various samples which help in batch selection for those image samples that the training models are most confused about. The selected batches are retrained with the same models for 10 active learning iterations which resulted in ~100% - 15% mIOU performance increase reflecting the promise of Trustworthy AI.</li></ul></li><li>□ <b>Data Analyst</b>, <a href="#">British Telecommunications Plc.</a>(worked as part of TCS), India [Oct, 2018 – Dec, 2019]<ul style="list-style-type: none"><li>• Primary responsibilities include to develop and maintain the State-of-Art Data Warehousing Portal Application for British Telecom and managing Innovative POCs for BT Wholesale' Cloudera Spark Setup.</li><li>• Member of the Decision Support System(DSS) Group to provide instant solutions to all British Telecom Business Verticals in accordance to AGILE Methodologies. Analyze customer habits, trends and feedbacks on new product-lines to provide a better insight to solve customer expectations and issues.</li><li>• Preparing and Presenting End-to-End Statistical OBIEE Report Generation and Visualizations with Tableau and Qlikview, D3.js.</li></ul></li><li>□ <b>ETL Developer</b>, <a href="#">British Telecom Plc.</a>(worked as part of TCS), India [Dec, 2016 – Sep, 2018]<ul style="list-style-type: none"><li>Part of the Data Insight(DI) team to support and address the largest telecom EDW migration in UK. About 85 TB of live database was seamlessly migrated from an Oracle 10g RAC to Oracle 12c RAC without losing any BAU critical timelines.</li></ul></li></ul>	
<b>Skills</b>	<ul style="list-style-type: none"><li>• <i>Programming Languages</i> : Proficient in Java, Scala, C++, ARM Assembly(ARMv7), CUDA(GPU) programming.</li><li>• <i>Scripting</i> : Shell, Scala, R, MATLAB, VHDL, WiMAX, Python.</li><li>• <i>Cloud Computing</i>: AWS Cloud, Microsoft Azure Cloud</li><li>• <i>ML Frameworks</i>: OpenCV, NLTK, Scikit-Learn, Pandas, NumPy, Matplotlib, mlpack, CARET</li><li>• <i>Deep Learning Models</i> : ImageNet(CNN) Pretrained, Seq2Seq(RNN), DCGAN (GAN), Im2Txt(CNN &amp; RNN), textsum(RNN), Large Language Models - various BERT, GPT versions .</li><li>• <i>Database</i> : SQL, MySQL, MongoDB, NoSQL Database(OrientDB).</li><li>• <i>Revision Control</i> : Git, SVN, Perforce.</li><li>• <i>Web Design</i>: PHP, HTML, CSS, BOOTSTRAP, Javascript, Jquery, AJAX, angular.js.</li><li>• <i>Apache Hadoop Ecosystem</i> : MapReduce, Spark, HDFS.</li><li>• <i>Data Format</i> : XML, JSON, CSV</li><li>• <i>Data Visualizations and Other Tools</i>: SAS 9.X, Ab Initio, Informatica PowerCenter, Oracle Data Integrator 12c, Tableau Desktop, D3.js, GGplot</li></ul>	
	<b>Expertise :</b> <ul style="list-style-type: none"><li>=&gt; Experience with various techniques in both Supervised and Unsupervised Learning like Regression, Clustering, Text Mining, Classifications, Decision Trees Etc.</li><li>=&gt; Strong working knowledge with machine learning algorithms such as k-NN, k-means, Naïve Bayes, SVM, Random Forest, Linear and Logistics Regression, Decision Tree(CART), Gradient Boosting and AdaBoost, Dimensionality Reduction Algorithms.</li><li>=&gt; Experience in building Deep Learning models like RNN(LSTM,GRU) and CNN(ResNet, InceptionV3,VGG16) and frameworks such as PyTorch, TensorFlow, Keras,Caffe, OpenCV<ul style="list-style-type: none"><li>- Developed Interactive Generative ChatBot with Seq2Seq RNN and RASA NLU</li><li>- Developed Speech Recognition Systems Text to Speech, Voice to Text</li><li>- Developed Image Captions(IR) using TF-Im2Txt for efficient Video and Image Search</li></ul></li></ul>	
<b>Publication</b>	<ol style="list-style-type: none"><li>1. D Goswami, S Chakraborty, <i>Active Learning for Video Classification with Frame Level Queries</i> , IEEE International Joint Conference on Neural Networks (IJCNN) 2023.</li><li>2. D Goswami, S Chakraborty, <i>Active Batch Sampling for Multi-label Classification with Binary User Feedback</i> , Winter Conference on Applications of Computer Vision (WACV) 2024 (Accepted)</li><li>3. K Roy, D Goswami 'Sentiment Analysis using Twitter Data for Demonetization in India', 5th International Conference on Business Analytics and Intelligence(BAICONF-17).</li></ol>	

**JESSICA MAYER**  
Raleigh, NC 27607  
(704) 993-1463 ♦ jkmayer2@ncsu.edu

## EDUCATION

**NORTH CAROLINA STATE UNIVERSITY**, Raleigh, NC

Expected 2025

**JOINT DEPARTMENT OF BIOMEDICAL AND HEALTH SCIENCE ENGINEERING, UNC-CHAPEL HILL & NC STATE UNIVERSITY** - B.S. in Biomedical and Health Science Engineering

## PROJECTS

### **CISCO 3D Printing and Design (TEAM MANAGER)**

Spring 2022

Conceived a design for a travel product utilizing proficiency with SOLIDWORKS and 3D printed functional models. Launched the product at an engineering design convention.

### **Engineering Grand Challenge Project (TEAM MANAGER)**

Fall 2021

Researched and promoted ways to continue efforts of reverse engineering the brain. Developed a creative solution to monitor sleep patterns and wake one during their lightest stage of sleep and improve sleep habits over time.

### **Sophomore Biomedical Design Project**

Fall 2022

Revitalized a prosthetic leg prototype utilizing AUTOCAD, Arduino coding and circuits, and hardware, creating a comfortable built-in cooling system to reduce sweat and discomfort.

### **Biological Pathogens Research, Cestoda**

Spring 2023

Pioneered and presented a research paper on tapeworms. Identified Cestoda history, lifecycle, mode of transmission, anatomy, host effects, treatments, and areas of public concern.

## LEADERSHIP & COMMUNITY INVOLVEMENT

### **NC State University Honors Program Member**

- Attend Scholar's Seminars led by valuable community members and world leaders
- Engage in community service
- Enroll in advanced honors and leadership courses

### **Helping Hand Club Member**

- Design and construct prosthetic limbs for children in need completely free of cost to the child and family utilizing SolidWorks, 3D printing, and engineering workshops

### **Chi Omega Fraternity Member and House Manager**

- Coordinate daily with house corporation board, house mom, executive board, and lawyers
- Manage meal plans and parlor fee finances for 130+ women
- Execute safety protocols in crisis events
- Plan and execute bonding events
- Assess and fix house safety, repair, quality, and food concerns

### **Biotechnology Club Member**

## WORK HISTORY

### **Spring Dermatology - Houston, TX**

June 2023 - August 2023

#### MEDICAL RECEPTIONIST AND ASSISTANT

- Assisted 20+ patients daily with prescriptions and accurately answered medical questions while providing empathetic patient care
- Proficiently managed electronic health records while maintaining HIPAA compliance, and developing data management expertise
- Achieved a 30% increase in clinic efficiency through strategic scheduling and organization
- Ensured accurate financial processing and attention to detail analyzing 50+ patient's insurance information per day
- Explored medications and needed medical advances through shadowing doctors and interacting with medical sales representatives

### **Piedmont Plastic Surgery And Dermatology - Charlotte, NC**

June 2022 - August 2022

#### MEDICAL DOCUMENT PROCESSOR

- Streamlined document processing system by digitizing and automating 15+ workflows; eliminated 2-year backlog of patient information with 99% accuracy, reducing processing time by 50%
- Digitized 5000+ physical documents using document management software, enabling easy access and reducing retrieval time by 30%
- Recovered 50+ corrupted documents through data recovery techniques, thereby reducing data loss by 10%

## TECHNICAL SKILLS

Research & Technical Report Writing Proficiency

Data Processing

Clinical Lab Safety Standards and Processes

Materials Testing

Microsoft Office Applications

Project Management

AutoCAD Proficiency

Engineering Drawings

Electronics & Circuit Design

Mammalian Cell Culturing

# Banmeet Kaur

kaurbanmeet100@gmail.com | 516-614-7223 | linkedin.com/in/banmeet-kaur- | github.com/banmeetkaur1

## EDUCATION

### Hofstra University

Bachelor of Science in Computer Science

December 2024

GPA: 3.75

- Programming Languages: Python, Java, C, C++, C#, Flutter/Dart, Perl, F#
- Tools/Libraries: Microsoft SQL, PostgreSQL, NumPy, Pandas, Scikit-learn, Matplotlib, Seaborn
- Secretary of Sikh Student Association

## TECHNICAL COURSEWORK

Machine Learning & Finance, Principles of Database Management, Programming Languages, Mobile Device Programming, Discrete Structures, Automata Theory, Computer Operating Systems & Architecture, Embedded Systems Design, Circuit Design Lab, Linear Algebra, General Physics, and Multivariable Calculus

## EXPERIENCE

### Research Assistant | Hofstra University

October 2023 – Present

- Collaborated on hardware and software development for robots aimed at assisting autistic children
- Utilized embedded systems to create advanced robots with features including audio, recording, mobility, and mobile app integration

## PROJECTS

### Data Science | Formula 1 Result Predictor

- Conducted in-depth analysis into historical and current Formula 1 data, including circuit details, weather conditions, team statistics, and driver performance
- Achieved an accuracy score of 84% with a machine learning model developed using Python, Pandas, and Scikit-learn

### Data Science | Credit Card Fraud Detection

- Developed a fraud detection machine learning model using Python, Pandas, and Scikit-learn, integrating exploratory data analysis (EDA) and achieving an accuracy score of 98%

### Embedded Systems | Smart Home Security System

- Engineered a comprehensive security system with multiple modes to protect homeowner property from intrusions throughout the day
- Utilized Arduino Uno R3, LEDs, speakers, and object detection sensors to monitor the environment for motion, activating different alarm modes & LED configurations based on the time of day

### Data Analysis | ETF Recommender

- Designed a Python web parser to obtain data on the largest ETFs and their holdings
- Built a recommendation engine of lower fee alternatives, thus improving investment decisions

### Data Analysis | Hedge Fund Trade Analysis

- Performed exploratory data analysis (EDA) of a trade fund in Python
- Generated insights and analytical recommendations using statistical tools and methods

## SKILLS

### Graphic Design/Video Editing

- Canva, CapCut, Adobe Premiere Pro

### Microsoft Applications

- Strong knowledge of Microsoft Excel, Word, and PowerPoint

### Multilingual

- English, Spanish, Punjabi, Hindi

# **Neel Patel**

Atlanta, GA | (501)-764-7976 | [neelhpatel2011@gmail.com](mailto:neelhpatel2011@gmail.com) | U.S. Citizen |  
<https://www.linkedin.com/in/neelhpatel2011/>

## **SUMMARY OF QUALIFICATIONS**

---

As a Biomedical Engineering student at Georgia Tech, my hands-on experience in engineering projects, laboratory research, and leadership roles have honed my abilities in problem-solving, communication, and project management, making me an ideal candidate for roles in healthcare.

**Python | C++ | MATLAB | Machine Learning | TensorFlow | Neural Networks | NLP | Cell Culture Techniques | Genomics | Bioinformatics | SolidWorks | LabView | 3D Printing | Microsoft Suite | Scientific Communication | Organizational Skills | Student Leadership | Data Analytical Skills | Project Management | Collaborative Teamwork**

## **EDUCATION**

---

**M.S. in Biomedical Engineering at Georgia Institute of Technology** August 2023 – Dec 2024  
**B.S. in Biomedical Engineering at University of Arkansas** August 2019 – May 2023  
Minor in Mathematics | GPA 4.0 | Honors College Fellow | Tau Beta Pi Member

## **ENGINEERING PROJECTS**

---

**Insertable Foot Sensor** August 2022 – August 2023

- Designed an insertable low-profile foot sensor capable of accurate force measurements in various locations
- Collaborated with 3 electrical engineers and 4 biomedical engineers to integrate electrical circuits within a patient's shoe
- Quantified walking asymmetry through custom MATLAB code force-time sensor measurements
- Enhanced patient comfortability by designing electrical components box on the shoelace rather than the ankle
- Awarded 2<sup>nd</sup> place at Engineering Exposition based on quality and engineering design

**Development and Analysis of Bone Cement** November 2021

- Designed high viscosity and high molecular weight of 350kDa bone cement of cylindrical shape using PMMA for hip replacement
- Operated 3400 series Instron compression testing to derive a multitude of mechanical properties such as Young's Modulus, ultimate tensile stress, and toughness using Microsoft Excel and MATLAB
- Calculated a Young's modulus of 59.32 MPa and tensile stress of 5.41, which is similar to industrial medical device properties for small bone constructs
- Presented bone cement design and analysis to UARK biomedical engineering faculty

**Research and Analysis of Synthetic Skin Substitutes** October 2021

- Determined two optimal synthetic skin substitutes from a variety of polymers through Instron tension testing
- Tested silicone skin model and calculated a Young's modulus of 748kPa with a percent error of 6.8%
- Concluded that silicone-derived materials exhibit skin-like properties through analysis of mechanical properties such as Young's modulus, ultimate strength, and percent elongation

**Node.js Tesla Car API** June 2020 – August 2020

- Coded and executed Tesla car commands, such as retrieving car speed, location, and battery conditions using Tesla's unofficial API
- Designed a functional, global-ranged, home garage opener operating from any computer by utilizing the unofficial Tesla API

## RESEARCH INTERNSHIP EXPERIENCE

### Lab Technologist/ Stem Cell Researcher

October 2020 – July 2023

- Analyzed mechanical stimulation in mesenchymal stem cells for improving osteogenesis through western blotting analysis of specific markers
- Published a co-authored literature-review manuscript on mesenchymal stem cell-based therapies titled: *Clinical utility of mesenchymal stem/stromal cells in regenerative medicine and cellular therapy*
- Skilled in mammalian cell culture, including subculturing, cryopreservation, and cell-thawing
- Demonstrated proficient protein analysis and purification utilizing techniques such as gel electrophoresis, Western Blotting, and Blot imaging using the ChemiDoc imaging systems to analyze cytoskeletal components in mesenchymal stem cells
- Utilized fluorescence microscopy for cell viability assays and mesenchymal stem cell imaging
- Executed techniques relating to RNA quantification, such as RNA isolation, reverse transcription, and qPCR for stem cell genome analysis
- Trained and managed four graduate students on stem cell culture, cell mechanical stimulation, Western blotting, and BCA protein assay

## PROFESSIONAL LEADERSHIP

### EMPOWER Peer Mentor – Univ. of Arkansas

August 2022 – Present

- Mentored freshmen college individuals with mild intellectual disabilities to further academic learning and foster inclusive socialization with a weekly commitment of 3 hours
- Served as a liaison between mentee's professors and EMPOWER supervisors to ensure mentee academic goals are met

### Resident Assistant – Univ. of Arkansas

August 2020 – Dec 2021

- Developed an inclusive community among a floor of 50 residents by coordinating monthly floor programs
- Worked 20 hours/week monitoring building safety, managing hall front desks, and leading weekly staff meetings
- Directed 80+ freshmen Lead Hogs students by fostering leadership development and inspiring creativity through monthly dorm programs and weekly 1-hour meetings
- Conducted 1-on-1 conversations with residents to examine adjustment to college life and determine needs by connecting them to on-campus resources and programs

## INTERNATIONAL EXPERIENCE

### Study Abroad in Sweden and Denmark – DIS

May 2022 – August 2022

- Researched drug development techniques and government regulations in Denmark biotech companies through various company tours and discussions with company representatives
- Investigated cellular and molecular neuroscience of various diseases by compiling brain tissue staining images at Karolinska Institute, Sweden using various brain databases such as The Human Protein Atlas and Allen Brain Atlas. Presented findings to Karolinska Neuroscience department

# CAN LI

📞 614-214-6126 📩 sapphirelican@gmail.com

## EDUCATION

<b>The University of Texas Health Science Center at Houston</b> Doctor of Philosophy, Biostatistics and Data Science	Houston, TX <i>Aug. 2020 - May 2025 (Anticipated)</i>
<b>Emory University</b> Master of Science in Public Health, Biostatistics	Atlanta, GA <i>Aug. 2018 - May 2020</i>
<b>The Ohio State University</b> <i>Dual degree</i> Bachelor of Science, Biochemistry Bachelor of Science, Applied Mathematics	Columbus, OH <i>Aug. 2013 - May 2018</i> <i>May 2016 - May 2018</i>

## PUBLICATIONS

1. **Li, C.**, Ding, S., Zou, N., Hu, X., Jiang, X., & Zhang, K. (2023). Multi-task learning with dynamic re-weighting to achieve fairness in healthcare predictive modeling. *Journal of Biomedical Informatics*, 143, 104399.
2. **Li, C.**, Jiang, X., & Zhang, K. (2023). A Transformer-Based Deep Learning Approach for Fairly Predicting Post-Liver Transplant Risk Factors. In arXiv [cs.LG]. arXiv. <http://arxiv.org/abs/2304.02780>. (Under Review in *Journal of Biomedical Informatics*)
3. **Li, C.**, Lai, D., Jiang, X., & Zhang, K. (2023). FERI: A Multitask-based Fairness Achieving Algorithm with Applications to Fair Organ Transplantation. In arXiv [cs.LG]. arXiv. <http://arxiv.org/abs/2310.13820>. (Submitted to American Medical Informatics Association (AMIA) 2024 Informatics Summit)
4. Upadhyaya, P., Zhang, K., **Li, C.**, Jiang, X., & Kim, Y. (2023). Scalable Causal Structure Learning: Scoping Review of Traditional and Deep Learning Algorithms and New Opportunities in Biomedicine. *JMIR Medical Informatics*, 11, e38266.

## CONFERENCE PRESENTATION

<b>International Conference on Intelligent Biology and Medicine (ICIBM 2023)</b> Presented “A Transformer-Based Deep Learning Approach for Fairly Predicting Post-Liver Transplant Risk Factors”.	<i>July 16-19, 2023</i>
------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-------------------------

## RESEARCH PROJECTS

<b>Equitable Rates in Multitask Learning for Fair Predictive Modeling</b>	<i>July 2023 - Present</i>
<ul style="list-style-type: none"><li>• Developed an algorithm Fairness through the Equitable Rate of Improvement in Multi-task Learning (FERI) to fairly predict graft failure risk in liver transplant patients by constraining each demographic subgroup's loss and equalizing improvement rates.</li><li>• Implemented the FERI algorithm in a large-scale liver transplant dataset, reducing the demographic parity disparity for gender by 71.74% and the equalized odds disparity for age groups by 40.46% while maintaining the prediction accuracy and enhancing the fairness of graft failure outcomes across demographics.</li><li>• Establishing a method using Multi-task learning techniques to improve fairness in clinical predictive models by equalizing training rates across tasks and patient subgroups.</li><li>• Improving the clinical predictive model's generalizability and robustness for imbalanced subgroup data through further validation and optimization.</li></ul>	

<b>Post-Liver Transplant Risk Predictions using Multi-Task Learning and Fairness Algorithms</b>	<i>Feb. 2023 - July 2023</i>
<ul style="list-style-type: none"><li>• Created a Transformer-based neural network for multi-task learning to simultaneously predict five crucial post-transplant risk factors with high accuracy performance across risks.</li><li>• Designed a task-balancing method and a fairness-achieving method to achieve a notable 39% reduction in task discrepancies and ensure unbiased predictions across demographic subpopulations.</li><li>• Structured a unified algorithm applying task-balancing and fairness-enhancing methods to improve donor-recipient matching in liver transplants and minimize disparities in attributes such as gender, age group, and race/ethnicity.</li></ul>	

<b>Fairness Enhancement in Sepsis Mortality Prediction using Dynamic Multi-Task Learning</b>	<i>May 2022 – Jan. 2023</i>
<ul style="list-style-type: none"><li>• Devised Multi-task learning with dynamic re-weighting to address bias in Sepsis mortality predictions in demographic groups by adjusting weights dynamically for multiple tasks, dividing predictions for diverse sub-populations, and approaching fairness as a task-balancing challenge.</li></ul>	

- Utilized the dynamic re-weighting method to achieve a 98% reduction in subgroup disparities and a 4% prediction accuracy sacrifice, strengthening more equitable and effective healthcare decision-making processes.

### Multi-Antihypertensive Drugs Individual Treatment Effects on Alzheimer's Disease

Oct. 2020 – Dec. 2021

- Derived causal structures integrating diagnosis, medical records, and socioeconomic features in the AthenaHealth Electronic Health Record (EHR) dataset.
- Emulated a target trial using the AthenaHealth EHR dataset to mimic the characteristics of a randomized trial.
- Performed propensity score matching and inverse propensity weighting to balance the differences in covariates, ensuring similar distributions of confounders between the treatment and control groups in Alzheimer's patients.
- Estimated treatment effects of multi-antihypertensive drugs for subgroups of patients using the Meta-Learners with feature importance interpretations.

## RESEARCH EXPERIENCE

---

### Graduate Research Assistant

Sept. 2022 - Aug. 2023

*Department of Bioinformatics and Computational Biology*

The University of Texas MD Anderson Cancer Center

Houston, TX

- Built an end-to-end pipeline for analyzing complex multi-modal Patient Derived Xenograft model dataset using deep learning models, including RNNs and feedforward networks through five-fold cross-validation and hyperparameter tuning, achieving high accuracy in predictions of drug treatment outcome.
- Conducted an in-depth spatial analysis of multiplex datasets utilizing UMAP-based visualization techniques to extract cancer-related genes.

### Graduate Research Assistant

Aug. 2020 - May 2022

*McWilliams School of Biomedical Informatics*

The University of Texas Health Science Center at Houston

Houston, TX

- Mitigated bias in AI systems and integrated deep learning models with multi-task strategies to enhance algorithmic fairness and accurate predictive modeling.
- Examined causal analysis on Alzheimer's Disease data with Meta-Learners to reveal heterogeneous treatment effects and disparities in the AthenaHealth EHR dataset.

### ORISE Fellow

Sept. 2019 - May 2020

Division of High-Consequence Pathogens and Pathology/Prion and Public Health Office (DHPP/PPHO), Centers for Disease Control and Prevention (CDC)

Atlanta, GA

- Specialized in epidemiological studies of infectious diseases, primarily focusing on Pasteurellosis hospitalizations and emergency department visits (2001-2017) in the United States.
- Employed SAS Macro programs and SQL to conduct statistical analyses on datasets such as HCUP's NIS and NEDS to learn disease trends over the years.
- Extracted Pasteurellosis-related data using ICD-9-CM and ICD-10-CM codes, emphasizing cases linked to dog and other animal bites. Analyzed demographic factors to determine hospitalization and ED visit trends for Pasteurellosis, subsequently calculating annual and average rates while identifying key demographic risk determinants.

## WORK EXPERIENCE

---

### Graduate Teaching Assistant

Sept. 2023 – Present

*School of Public Health*

The University of Texas Health Science Center at Houston

Houston, TX

- Course:** Categorical Data Analysis (Master & PhD Level)

### Data Science Intern

May 2022 – Aug. 2022

Data Science Analytics & Insights, Real-World Evidence for Janssen Research & Development, Johnson & Johnson

Remote

- Developed a robust cohort construction in Optum EHR to characterize patient profiles and quantify risk factors, creating the events and censored patient identification.
- Employed the Charlson comorbidity index (CCI) for risk stratification and designed a Cox regression model, integrating demographic data, comorbidities, hospitalization duration, and disease stage.
- Interpreted hazard ratios using forest plots and assessed disease recurrence potential, providing decision-makers with critical insights for potential inclusion/exclusion criteria in precision medicine.

## TECHNICAL HIGHLIGHTS

---

**Programming Languages:** Python, R, SAS, Java, SQL, and MATLAB.

**Expertise:** Machine Learning and Model Development, Algorithms, Deep Learning, Bias mitigation and Fairness, Predictive Modeling, Data Visualization, and Regression Analysis.

# Austin Rush

## Software Developer

238 E Broad St • Souderton, PA 18964 • [AR1025019@wcupa.edu](mailto:AR1025019@wcupa.edu) • 267-977-5100 • <https://github.com/austincrush>

---

### Education

#### West Chester University

Computer Science 3.4

West Chester, PA

August 2023 – Current

#### Montgomery County Community College

Computer Science 3.3 GPA

Blue Bell, PA

August 2021 – May 2023

#### Souderton Area High School

3.3 GPA

Souderton, PA

Graduated June 2021

---

### Projects

#### Snake Game

- Java program designed as the classic snake game.
- Swing library used to create the interactive GUI interface.
- Strong elements of Object-Oriented Programming and version control present.

#### Hangman Game:

- Designed an interactive game created in Java with the JavaFX library.
- Version control (GIT) used to keep track and manage updates.

#### Weather App:

- JavaFX app designed to give the user current weather information on the desired city.
- Interfaces with the WeatherStack.com API to get real time data.

---

### Skills

#### Programming Languages:

- Proficient: Java (Strong OOP Design), HTML/CSS
- Familiar: JavaScript, SQL, Python

**Software:** MySQL, Strong OOP Version Control (GIT/GitHub), IntelliJ, Windows, MacOS, Linux (Ubuntu), Excel, VSCode

**Soft Skills:** Leadership, Critical Thinking, Analytical, Communication

---

### Work Experience

#### Giant Food Stores

##### Grocery Associate

- Stock and organize shelves with product.
- Keep track of and adjust inventory counts.
- Unload and organize deliveries.

Souderton, PA

September 2018 – August 2023

##### Diary Associate

- Stock and organize shelves with product.
- Rotate perishable stock

August 2023 – Current



# Elizabeth Tsang

[etsang2016@gmail.com](mailto:etsang2016@gmail.com) □ (973) 727-2328 □ Mountain Lakes, NJ

---

## WORK EXPERIENCE

---

### Rutgers University - New Brunswick

**Jan. 2023 - May 2023**

*Data Structures Learning Assistant*

*Piscataway, NJ*

- Independently led three 55-minute sessions applying lecture skills to problems relating to data structures
- Strengthened student understanding by encouraging group discussion and collaboration
- Analyzed classroom environment and adapted quickly to discussion flow

### Mountain Lakes Swim & Dive Association

**Jul.**

**2021 - Aug. 2021**

*Assistant Swim Coach*

*Mountain Lakes, NJ*

- Demonstrating skill in the sport and effective teaching of techniques
- Encouraging and supporting swimmers to gain confidence and self-esteem
- Managing swimmers at competitions and making sure they get to events on time

### Lakeland Hills YMCA

**Mar. 2020 -**

**Aug. 2023**

*Lifeguard*

*Lakes, NJ*

*Mountain*

- Staying attentive to patrons swimming in the pool

- Being responsive and open to interacting with patrons

- Maintaining pool deck, water checks, closing facilities

## EDUCATION

---

### Rutgers University - New Brunswick

**May 2026**

*BS, Computer Science*

*Brunswick, NJ*

*New*

- Dean's List 3.571/4.0 GPA

- Member of Douglass SAS DIMACS Computer Science LLC
  - Living-learning community of women with an interest in computer science

- Member of Creation of Games Society

- o Discuss game theory/design, worked on development of games in groups
- Relevant coursework: Data Structures, Computer Architecture, Calculus I & II, Discrete Structures I, Principles of Programming Languages

## PERSONAL PROJECTS

---

### Piano

**May 2022**

*Java, Eclipse IDE*

- Developed from scratch as a final project for AP Computer Science A
- Displays piano keys in single octave range, responds to both keyboard and mouse clicks
- Usage of two objects: Instrument and Key

## SKILLS

---

- **Hard skills:** Java; Python; C; Linux; Assembly; learning OCaml in classroom
- **Soft skills:** Peer leadership; collaboration; teamwork; creativity; writing; communication; empathy

# Jordan Jomsky

89 Prospect Park SW, Brooklyn NY

805.208.2122 | jordan.jomsky@columbia.edu | linkedin.com/in/jordanjomsky

## EDUCATION

### Columbia University

*Master of Science in Data Science*

New York, NY

Dec 2024

Relevant course work: Algorithms for Data Science, Data Visualization, Natural Language Processing

### University of California, Berkeley

*Bachelor of Arts in Data Science with Honors and in Molecular and Cell Biology*

Berkeley, CA

May 2021

Courses: Principles of Data Science, Data Inference, Data Mining, Data Structures, Anatomy and Physiology

Honors: Coding it Forward Fellowship (Summer 2021), RHA President of the Year (Feb 2020)

Thesis: "Abnormal Nodule Detection on Chest X-Rays using Machine Learning"

## TECHNICAL SKILLS

**Python:** Data Analysis, Machine Learning, Artificial Intelligence, 7 years of experience

**Java:** Object-Oriented Programming, Unit Testing, Data Structures, Runtime Analysis, 8 years of experience

**SQL:** MySQL, Automated Data Extraction, Relational Databases, 4 years of experience

**R:** Statistical Analysis, Data Analysis, PDF Text Extraction, 4 years of experience

## WORK EXPERIENCE

### National Institute of Health

*Data Scientist*

Washington, DC

Aug 2021 - Present

- Managed institute-wide initiatives for automating SQL and R based workflows for grant auto encoding with recent promotion into the Office of Director within the Division of Data Science Strategy
- Designed information extractor with named entity recognition and Keras neural network architecture, decreasing categorizing process time by 99% for grants utilizing stem cells in research
- Started a new project to integrate ChatGPT using Azure into nationwide extramural grant applicants' search for institutional and departmental support benefiting thousands of applicants unfamiliar with NIH's process

### Coding it Forward Data Science Fellow

May 2021 - Aug 2021

- Initialized creation of a custom pipeline to automate grant classification using spaCy and Gensim in Python
- Standardized numerical representation of scientific papers with custom BERT transformers for an easy-to-use topic model that has identified three areas of scientific research with no direct counterpart in the department

### Mishler Lab

Berkeley, CA

*Research Apprentice*

Sep 2018 - May 2021

- Created a custom data visualization deploying Python and Tableau for spatial representation of genetic sex ratio data for moss samples to determine further experimentation and understanding
- Produced public coding repository with 2 datasets and full Jupyter notebook showcasing advanced statistical analysis methods utilized in project in name of open science
- Led contributions to methods and results section of peer-reviewed publication in Applications in Plant Sciences journal of the Botanical Society of America with 3 additional custom visualizations

### MiiCare

London, United Kingdom

May 2020 - Aug 2020

*Data Analytics and Software Engineering Intern*

- Collaborated with a team of global industry professionals in a \$4 million startup based in London creating a proprietary product to care for elderly leveraging artificial intelligence
- Developed a unique data structure in Java to perform time-based Markov chain analysis on routine of seniors using only data gathered by sensors throughout initial 10 users' homes
- Spearheaded creation of novel algorithms in Python using machine learning data generation to determine sleep patterns of seniors without use of any wearable tech

## PUBLICATIONS

Ekwealor, J. T. B., Benjamin, S. D., **Jomsky, J. Z.**, Bowker, M. A., Stark, L. R., McLetchie, D. N., Mishler, B. D., & Fisher, K. M. (2022). Genotypic confirmation of a biased phenotypic sex ratio in a dryland moss using restriction fragment length polymorphisms. *Applications in plant sciences*, 10(2), e11467. <https://doi.org/10.1002/aps3.11467>

# Hang Yu

Email: yuooo408@umn.edu | Phone: (508) 864 5856 | LinkedIn: hangyu3 | Google Scholar

## Summary

- Experienced in machine learning algorithms and multimodel time series data.
- Proficient in Python | Designed a publicly available Android App for sensor data streaming.
- Skilled in Linux environments and Github for code version control.

## Education

Ph.D.	University of Minnesota - Twin Cities	Bioinformatics and Computational Biology	Dec 2024
M.Sc.	University of Missouri - Columbia	Data Science & Analytics	Sept 2019
M.Sc.	Worcester Polytechnic Institute	Information Technology	Aug 2017
B.Sc.	Dalian Univ. of Foreign Lang.	Interpretation	July 2015

## Work Experience: Machine Learning (ML), Software (SW) Dev., Databases (DB)

University of Minnesota - Twin Cities	Research Assistant	Sept 2020 – Present
---------------------------------------	--------------------	---------------------

ML	Machine Learning for human activity prediction	Python	Tensorflow	Keras	Scipy	Git
----	------------------------------------------------	--------	------------	-------	-------	-----

- Processed sensor data (ECG, PPG, accelerometer) with filters and augmentation.
- Analyzed wearable sensor data (ECG, PPG) with signal processing.
- Developed machine learning models including CNN, LSTM to predict human activity. **Publication:** [PSB](#)
- Deployed pre-trained large models including wav2vec, Resnet and transfer learning for prediction.

SW	Android application for sensor data collection	Java	Android Studio	SQL	wearables
----	------------------------------------------------	------	----------------	-----	-----------

- Designed an integrated Android application for sensor and survey data streaming and storage.
- Developed the publicly available android app ([App Link](#)) for Polar® sensors for internal/external users.
- Wrote extensible codes with Java and SQL by following best SW dev practices.

ML	Image in-painting and Self-supervised Learning	Tensorflow	OpenCV
----	------------------------------------------------	------------	--------

- Leveraged transfer learning to diagnose arthritis on medical images.

University of Missouri - Columbia	Research Assistant	May 2019 – Sept 2020
-----------------------------------	--------------------	----------------------

DB	Hierarchical Time Indexed Sensor Data Repository Design	Python	SQL	MongoDB	R	Linux
----	---------------------------------------------------------	--------	-----	---------	---	-------

- Designed a hierarchical data repository by applying column-based arrayized structure. **Publication:** [IEEE](#)
- Ran experiments and designed hierarchical repository to optimize accessibility and stability of the system.

Home Source Industries	Data Analyst	Feb 2018 – Dec 2018
------------------------	--------------	---------------------

DS	Data analytics and visualization for purchasing and inventory	Tableau
----	---------------------------------------------------------------	---------

- Built data forecasting model for sales and inventory management.

NuVant Systems Inc. (Northeastern University)	Database Development Intern	Aug 2017 – Nov 2017
-----------------------------------------------	-----------------------------	---------------------

DB	SQL database design for battery recondition data	MySQL	MS Access
----	--------------------------------------------------	-------	-----------

- Developed MySQL database and used stored procedures to analyze data.

Worcester Polytechnic Institute	M.Sc. Projects	Aug 2015 – June 2017
---------------------------------	----------------	----------------------

DB	Database Application Development project	MySQL	Android Studio	Java
----	------------------------------------------	-------	----------------	------

- Designed an Android application to create an inventory and sales management system.
- Created a MySQL database implementing entity-relationship diagrams, UI design, and queries.

UI	Website User Experience Design for Parents	HTML	Tobii
----	--------------------------------------------	------	-------

- Utilized data analysis to find first mouse click, first fixation duration, and gaze point heat-maps.

# Roshan Patel

E-mail: [roshanp@princeton.edu](mailto:roshanp@princeton.edu) | Cell: 631-902-0225

## Education:

*Princeton University, 2020 – Present*

PhD Candidate, Chemical and Biological Engineering  
M.S Chemical and Biological Engineering  
Research Advisor: Michael A. Webb

*University of Pennsylvania, 2016 – 2020*

Vagelos Integrated Program in Energy Research (VIPER)  
B.A in Physics, B.S.E in Chemical and Biomolecular Engineering, *summa cum laude* (GPA 3.85)  
Undergraduate Research Advisor: Aleksandra Vojvodic

## Research Experience:

**PhD Candidate**, Princeton University

8/2020-Present

My research is focused on developing efficient computational methods used to design synthetic polymers that augment and/or mimic protein functions

- Trained and leveraged statistical ML models in an active learning workflow to design polymer excipient formulations that provide proteins with unprecedented stability against thermal stresses
- Closely collaborate with experimentalists to validate simulations, fine-tune molecular models, and guide high-throughput assays for rapid discovery using data-driven algorithms
- Developed and systematically investigated deep learning strategies (e.g., CNNs, LSTMs, GNNs) to predict physical properties of copolymers using experimental and simulated datasets
- Investigating the use of unsupervised representation learning and Bayesian neural networks for active learning strategies aimed at building predictive, sequence-to-property models
- Developed an algorithm to efficiently train coarse-grained, neural network potentials; using free energy perturbation and related methods to assess algorithm performance
- Performed quantum chemical calculations to parameterize new molecular dynamics models of bio-interfacing polymers; using this model to probe polymer – protein interactions
- Used unsupervised machine learning and molecular dynamics simulations to characterize the relationship between primary chemical structure and morphology in bio-mimetic polymers

**Undergraduate Researcher**, University of Pennsylvania

5/2017-5/2020

- Characterized the crystal structure of perovskite materials with applications in clean hydrogen fuel production using quantum chemical (QC) calculations and numerical optimization methods
- Performed QC calculations and used kinetic models to investigate how doping perovskite catalysts with rare-earth metals can enhance activity
- Performed QC calculations and developed thermodynamic models to study the stability of catalysts in electrolytic cells under operating conditions

## Peer Reviewed Publications:

**Patel, R. A.**, Colmenares S, Webb M. A., “Sequence patterning, morphology, and dispersity in single-chain nanoparticles: Insights from Simulation and Machine Learning.” *ACS Polymers Au*, 2023, <https://doi.org/10.1021/acspolymersau.3c00007>.

**Patel, R. A.**, and Webb M. A., “Data-Driven Design of Polymer-Based Biomaterials: High-Throughput Simulation, Experimentation, and Machine Learning.” *ACS Applied Bio Materials*, 2023, <https://doi.org/10.1021/acsabm.2c00962>.

Tamasi, M. J.\* , **Patel, R. A.\***, Borca, C. H.\* , Kosuri, S.\* , Mugnier, H., Upadhyay, R., Murthy, N. S., Webb, M. A., Gormley, A. J. “Machine Learning on a Robotic Platform for the Design of Polymer-Protein Hybrids.” *Advanced Materials*, 2022, <https://doi.org/10.1002/adma.202201809>.

**Patel, R. A.**, Borca, C. H., Webb, M. A., “Featurization Strategies for Polymer Sequence or Composition Design by Machine Learning.” *Molecular Systems Design & Engineering*, 2022, <https://doi.org/10.1039/d1me00160d>.

Kosuri, S., Borca, C. H., Mugnier, H., Tamasi, M. J., **Patel, R. A.**, Perez, I., Kumar, S., Finkel, Z., Schloss, R., Cai, L., Yarmush, M. L., Webb, M. A., Gormley, A. J. “Machine-Assisted Discovery of Chondroitinase ABC Complexes toward Sustained Neural Regeneration.” *Advanced Healthcare Materials*, 2022, <https://doi.org/10.1002/adhm.202102101>.

Raman, A. S., **Patel, R. A.**, Vojvodic, A., “Surface Stability of Perovskite Oxides under OER Operating Conditions: A First Principles Approach.” *Faraday Discussions*, 2021, <https://doi.org/10.1039/c9fd00146h>.

## **Manuscripts under preparation:**

**Patel, R. A.**, Webb, M.A., “Relative Entropy Minimization via Stochastic Gradient Descent: An Efficient Strategy for Machine Learning Coarse-Grained Potentials”.

**Patel, R. A.**, Webb, M.A., “Deep Bayesian Active Learning for Modelling Sequence-Dependent Polymer Properties: A Data-Limited Approach”.

## **Presentations and Posters:**

“Machine Guided Design of Polymer-Protein Hybrid Materials”. Foundations of Molecular Modeling and Simulation: Molecular modeling and the Data Revolution. 2022. Poster

“Constructing supervised machine learning models for copolymer property prediction: Featurization strategies.” American Society of Chemistry Middle Atlantic Regional Meeting 2022. Oral Presentation.

## **Skills:**

**Python:** numpy, scipy, pandas, scikit-learn, tensorflow, spectral, pytorch, jax, hyperopt, joblib, matplotlib, rdkit, mordred

**Molecular Dynamics:** LAMMPS

**Electronic structure / QC calculations:** ORCA, Quantum ESPRESSO  
**bash:** awk, sed  
**c/c++**  
**git/github**

## **Notable Awards:**

### American Chemical Society Excellence in Graduate Polymer Research (2023)

Recognized nationally in a group of 54 students for making impactful contributions to polymer science and engineering by leaders in the field. Invited to present work in a special symposium for award recipients in the upcoming ACS Spring 2024 National Meeting.

### Foundations of Molecular Modeling and Simulation Best Poster Presentation (2022)

One of five recipients of the “Best Poster” Award in a group of 101 presenters made up of graduate students, industry representatives, post-doctoral scholars, and university professors.

### Graduate Student Symposium Best Poster Presentation (2022)

Awarded “Best Poster Presentation” by faculty at Princeton University Department of Chemical and Biological Engineering amongst all 3<sup>rd</sup> year graduate students.

### National Science Foundation Graduate Research Fellowship Honorable Mention (2021)

“The NSF accords Honorable Mention to meritorious applicants who do not receive Fellowship awards. This is considered a significant national academic achievement and provides access to cyberinfrastructure resources through the XSEDE.”

### The Hugo Otto Wolf Memorial Prize (2020)

“The Hugo Otto Wolf Memorial Prize, founded by Otto C. Wolf in memory of his son, is awarded to that member of the senior class in each department of the School of Engineering and Applied Science who, during the senior year, by the thoroughness and originality of his or her work, meets with the greatest approval of the professors in charge.” Obtained as a senior award at the University of Pennsylvania for the department of Chemical and Biomolecular Engineering.

## **Teaching Experience:**

### CBE 512: Machine Learning in the Chemical Sciences (Princeton University)

Assistant Instructor: Created and graded problem sets, held office hours and extra help sessions

### CBE 231: Thermodynamics of Fluids (University of Pennsylvania)

Teaching Assistant: Created and graded problem sets and exam questions, held office hours and extra help sessions

# DOLLY RIJWANI

dollyrijwani22@gmail.com | linkedin.com/in/dolly-rijwani | (213) 619-6925

## EDUCATION

<b>Master of Science - Applied Data Science</b> <i>University of Southern California</i>	<b>August 2023 - May 2025</b> <i>Los Angeles, CA</i>
<b>Bachelor of Technology in Computer Science - Gold medalist</b> <i>Madan Mohan Malaviya University of Technology, Gorakhpur, India</i>	<b>June 2013 - July 2017</b> <i>GPA: 3.9/4</i>

## EXPERIENCE

<b>Software Engineer</b> <i>Vimeo Technologies</i>	<b>June 2021 - July 2023</b> <i>Bangalore, India</i>
<ul style="list-style-type: none"><li><b>Optimized User Registration Process</b>, streamlined queuing system using Beanstalkd and Google pub/sub reducing registration processing time by <b>50%</b>, resulting in <b>10%</b> increase in webinar attendance</li><li><b>Revenue-Boosting Analytics</b>, developed an algorithm to generate viewer analytic reports, integrating real-time email editing templates, leading to a <b>\$100,000</b> increase in revenue</li><li><b>Seamless Third-Party Integration</b>, integrated third-party <b>email service providers</b> like Marketo, Hubspot, and Constant Contact, simplifying the registration process, resulting in a <b>15%</b> increase in user sign-ups.</li></ul>	<b>July 2017 – May 2021</b> <i>Bangalore, India</i>

## PROJECTS

<b>Credit risk model using Logistic Regression</b>	<b>February 2022 - March 2022</b>
<ul style="list-style-type: none"><li><b>Problem:</b> Building a binary loan defaulter classification model for credit risk assessment</li><li><b>Approach:</b> Employed CRISP-DM framework, including exploratory Data Analysis, standardization, <b>log likelihood cost function</b>, and gradient descent with Python and Numpy, utilized <b>chi-square test</b> for goodness of fit</li><li><b>Result:</b> Achieved an <b>80%</b> recall rate, producing a final credit risk model using Logistic Regression, leading to <b>reduced financial risk</b> for the company and more informed lending decisions</li></ul>	
<b>Dementia prediction using Support Vector Machine</b>	<b>December 2021 - January 2022</b>
<ul style="list-style-type: none"><li><b>Problem:</b> Detection of mild cognitive impairment (MCI) using SVM with OASIS Longitudinal brain MRI Dataset</li><li><b>Approach:</b> Implemented cross-validation, <b>grid search</b> for hyperparameter tuning, and AUC analysis for threshold selection, optimizing <b>bias-variance trade-off</b></li><li><b>Result:</b> Attained a <b>70%</b> recall rate in test data for improved MCI prediction, enhanced <b>early detection</b> of mild cognitive impairment (MCI), potentially reducing healthcare costs</li></ul>	
<b>Face Mask Detection using Convolutional Neural Network</b>	<b>November 2021 - December 2021</b>
<ul style="list-style-type: none"><li><b>Problem:</b> Classify images into masked, unmasked, or partially masked categories</li><li><b>Approach:</b> Applied the LeNet-5 architecture, Adam Optimizer, and image augmentation, conducted parameter search for optimization, exploring <b>activation function</b>, alpha, beta, and neuron count</li><li><b>Result:</b> Obtained a <b>90%</b> accuracy rate on the test set and <b>93%</b> accuracy on the validation set, <b>improved public safety</b> and awareness, leading to greater adherence to mask-wearing guidelines.</li></ul>	

## TECHNICAL SKILLS

**Programming Languages:** Python (Numpy, Pandas), SQL, NoSQL, PHP, Javascript

**Web Development:** HTML5, Django, React, Angular, Git, Docker

**Data Science :** Tensorflow, Scikit-Learn, PyTorch, Hadoop, Spark, Matplotlib, Tableau, Hypothesis testing

# Tanay Dixit

EDUCATION [✉ tanayd2@illinois.edu](mailto:tanayd2@illinois.edu) [\(217\) 305-1537](tel:(217)305-1537) [tanay](#) [GScholar](#) [tanay2001.github.io](#)

University of Illinois Urbana Champaign

Aug 2023 - May 2025

MS in Computer Science

Indian Institute of Technology (IIT), Madras, India

July 2023

B.Tech in Electrical Engineering and Minor in Computing

## WORK EXPERIENCE

Microsoft Research | Research Intern

March - Jun 2023

Code Generation using Large Language Models

- Conducted benchmarking on the SPIDER dataset using **LangChain** to assess the text-to-SQL capabilities of **LLMs**.
- Increased the control over LLMs for text-to-SQL generation by introducing intermediate representations.
- Integrated our approach into **Azure Data's Trident** project, boosting engagement by 38.3% based on a user study.

Subex AI Labs | Machine Learning Intern

Jun - Aug 2021

Automating document verification

- Curated an in-house invoice dataset to benchmark model performance and quantify improvements.
- Designed an OCR pipeline using **Tesseract** to extract text from structured and unstructured invoices.
- Implemented a document layout recognition model by fine-tuning the **LayoutLM-v2** model on the custom dataset.

BlueBarrel Solutions Pvt Ltd | Data Scientist Intern

Nov - Dec 2021

Analyzing and estimating energy usage

- Analyzed 10 million+ energy consumption data points using **SQL** & mined patterns to recommend power suppliers.
- Developed an **interactive dashboard** using flask & plotly to reduce exploratory data analysis time by 40%.

## RESEARCH EXPERIENCE

Research Assistant | University of Illinois Urbana-Champaign

Aug 2023 - Current

Fact Checking claims in the era of Large Language Models

- Incorporating LLMs with symbolic solvers for interpretable and accurate fact-checking of complex claims.

Bachelor Thesis | IIT Madras [Publications: [ACL 2023](#), [EMNLP 2021](#)]

Jan 2021 - May 2023

Towards Analysing and Improving Natural Language Generation (NLG) evaluation metrics

- Analyzed the correlation of evaluation metrics with human evaluation criteria using **SciPy** and **NumPy**.
- Designed **perturbations** to highlight some pitfalls in current NLG metrics and to aid in better design of metrics.
- Analyzed **evaluation metrics** and optimized the **COMET** metric using Pytorch for evaluation in Indic languages.

Research Intern | University of Southern California [Publications: [ACL 2023](#)]

June - Oct 2022

Improving Factuality of Abstractive Summarization without Sacrificing Summary Quality

- Proposed a ranking strategy by combining different metrics for contrastive fine-tuning **BART** and **PEGASUS**.
- Demonstrated that our approach showed up to 18% gains on factuality metrics without sacrificing abstractiveness.

Research Intern | University of Washington [Publications: [EMNLP 2022](#)]

Jan-May 2022

Counterfactual data generation for model interpretability

- Fine-tuned DPR for developing a **retrieval-augmented** framework for generating diverse counterfactuals.
- Evaluated the effectiveness of our approach in improving RoBERTa and DeBERTa's **robustness on NLU tasks**.

## KEY PROJECTS

Multilingual Text Classification and Headline Generation |



- Designed a unique phoneme-based kim-CNN architecture using **PyTorch** for text classification. March 2021
- Fine-tuned **mBART** model using **Transformers** for improved accuracy on low resource languages like *Hindi*.

Analysis of Recommendation System |



- Optimized the performance of **matrix factorization** on ranking metrics like MAP, nDCG by 22% July 2020
- Used novelty metrics to analyze the ability of systems to recommend items from the long tail.

## TECHNICAL SKILLS

- **Languages:** Python, C++, C, SQL, Matlab, Bash
- **Libraries/Tools:** PyTorch, Tensorflow, LangChain, Docker, Transformers, Scikit-Learn, NumPy, SciPy, Pandas, Git

# Ross Chu

Machine Learning Scientist (Ph.D.)

	<b>E-mail:</b>	ross.chu@berkeley.edu
	<b>Phone:</b>	+1 (510) 495-0634
	<b>Location:</b>	San Francisco Bay Area (Berkeley, CA)
	<b>Profile:</b>	<a href="https://www.linkedin.com/in/rosschu">https://www.linkedin.com/in/rosschu</a>

## Education

### University of California, Berkeley

Ph.D. Economics

2019 – 2025

(Expected)

- EXPERTISE: Labor Markets & Mathematical Psychology
- DISSERTATION: Behavioral Models of Salary Negotiations
- AWARDS: NSF Research Fellowship (\$145,000), DCF Dissertation Fellowship (\$36,000), Strandberg Research Award (\$4,000)

### Northwestern University

Graduated 2017

B.A. Mathematics, B.A. Statistics, B.A. Economics

- THESIS: Causal Impact of Home Country Disasters on Immigrants' Exam Performance
- AWARDS: Summa Cum Laude (GPA: 3.97/4.00), Phi Beta Kappa (< 10%), Deibler Honors Thesis Award ( $\leq 3$  students)

## Experience

### Korean Statistics Data Center & UC Berkeley

Aug 2021 – Current

Dissertation Researcher (Economics, Ph.D.)

Seoul, Korea / Berkeley, CA

- Examining how cognitive biases affect salary negotiations by developing mathematical models and estimating them with salary data
- Analyzing two-sided labor markets by solving dynamic optimization problems with recursive algorithms from reinforcement learning
- Quantifying salary impacts of labor policies in Europe by conducting computational simulations with the estimated model in Python
- Improving estimates of gender pay gaps by flexibly controlling for workplace benefits with double-debiased machine learning
- Designing experiments to measure changes in negotiation behavior when AI predictions are used as heuristics for salary comparisons

### Amazon.com (Science Team @ Workforce Planning Solutions)

Jun 2022 – Aug 2022

Economist / Data Scientist (Ph.D. Intern)

Seattle, WA

- Spearheaded the first impact evaluation of a workforce training program that affects 23% of Amazon associates at fulfillment centers
- Adapted estimates of program impacts to each fulfillment center by combining causal inference with Bayesian Shrinkage methods
- Compiled business reports with data visualizations of performance metrics by processing trillions of data points with SQL and Python
- Advised on strategic improvements for the training program by presenting key findings to organization directors and product managers

### Department of Economics (UC Berkeley)

Dec 2019 – Dec 2020

Graduate Student Researcher

Berkeley, CA

- Automated a large-scale experiment for detecting discriminatory hiring practices for entry-level jobs at Fortune 500 companies
- Saved 11,000+ hours of manual labor by developing software that scrapes online job postings and automatically submits applications
- Constructed a data-centric measure of resume quality by leveraging machine learning to predict the probability of interview callbacks

### National Bureau of Economic Research (Harvard University)

Jul 2017 – Jun 2019

Pre-Doctoral Research Associate

Cambridge, MA

- Evaluated the causal impact of auto-enrollment into 401(k) savings plans on employees' financial health using applied econometrics
- Forecasted the accumulation of retirement assets for auto-enrollees based on hypothetical stock returns experienced by non-enrollees
- Accelerated bootstrap algorithms with computational shortcuts and parallel computing on the Unix Shell, resulting in a 37x speed boost
- Disseminated research findings through data visualizations in financial reports, harnessing over 300 GB of data on 401(k) savings plans

### Institute for Policy Research (Northwestern University)

Apr 2016 – Jun 2017

Thesis Researcher (Economics, Honors B.A.)

Evanston, IL

- Authored an award-winning thesis that studies the causal relationship between immigrants' test scores and home country disasters
- Assembled a research dataset by scraping global disaster records with Python and linking them with test score data from Florida
- Estimated causal effects of unexpected disaster shocks by integrating them into a quasi-experimental framework for robust inference

## Skills & Leadership

- PROGRAMMING: Python, SQL, R, Stata, Unix Shell
- SKILLS: GitHub, Scikit-learn, Pandas, NumPy, SciPy, Seaborn/Matplotlib, Scrapy, Selenium, XPath, AWS, Google Cloud
- LEADERSHIP: Course Instructor (Applied Causal Inference), PhD Mentor for CS/Econ Undergraduates (Diversity + RA Programs), Squad Leader (Embassy Security Unit: Korean Police), Elected Captain of Trainee Cohort (Korean Police)

# Karthikeya Telapalli

Tampa, FL | 813-894-2873 | [kartikeyatelapalli@usf.edu](mailto:kartikeyatelapalli@usf.edu) | [LinkedIn](#) | Willing to Relocate

## EDUCATION

University of South Florida

Tampa, FL

**BS in Computer Science**

May 2025

- **Dean's list** for 2 consecutive semesters.
- Recipient of USF Green and Gold Directors award.
- Cumulative GPA: **3.75/4.0**

## WORK EXPERIENCE

University of South Florida

Tampa, FL

**Research Assistant @ Tu's Lab.**

09/2022 - current

- Employed **Agile methodologies** to Orchestrate the development of a **neural network algorithm** by leveraging convolutional neural networks(CNNs) to analyse intricate Gut **microbiome data**, culminating in the successful **prediction** of new data with an impressive **90% accuracy rate**. Optimised the model's performance by implementing regularisation techniques to improve its prediction power.
- Spearheaded a high-calibre team in the development of a cutting-edge 3D **Molecular Dynamics simulator** utilising **CUDA technology**. Designed and implemented the Verlet integration method to accurately model the dynamic motion of N particles within a 3D system.
- **Integrating** molecular dynamics simulations with **Apache Spark** for distributed processing and leveraging **JCuda** for GPU acceleration, encapsulated as a **Spring Boot-based web service**.

University of South Florida

Tampa, FL

*Undergraduate Research Society.*

12/2022 - 04/2023

- **Collaborated** with a team of Six members to conduct a **comprehensive analysis** of a vast dataset consisting of over 1000 articles, meticulously collecting crucial data and statistics related to BRCA 1 and BRCA 2 mutations and their relation to breast cancer(**Epidemiological Overview**). Presented key findings based on the data collected to a broad audience.

## PROJECTS

Automated **Malware Detection System** (March 2023)

- Designed and implemented **deep learning algorithm** to accurately **detect** malicious software culminating in an **accuracy rate of 95%**.

Social Media **Sentiment Analysis** (January 2023)

- Created a deep learning based sentiment analysis tool using Python, **TensorFlow**, and **Natural Language Processing** (NLP) techniques to analyse **social media data**, identifying trends and **insights** for **marketing strategies**. Performed Preprocessing on large datasets to ensure data quality and reliability in sentiment analysis.

Advanced Stock **Price Prediction** Tool (October 2022)

- Leveraged a multi-faceted approach to stock price forecasting by **integrating random forests with neural networks**. Achieved a remarkable prediction **accuracy of 80%**, further backed by comprehensive metrics like **MSE and MAPE**.
- Subjected the model to **rigorous validation**, including stress tests simulating major market upheavals, ensuring **reliability** and resilience across **diverse market conditions**.

## SKILLS & INTERESTS

Technical Skills: **Python, C, C++, Javascript Data Analytics, SQL, Natural Language Processing, Machine Learning, Jupyter Notebook, Tensorflow, NumPy, Scikit-learn, Algorithms, Statistical Modelling, Applied Mathematics, C#, Proficient in MS Office Suite, Pandas.**

Soft Skills: Analytical, Innovative, **Communication, Problem Solving, Emotional Intelligence, Conflict Management.**

## LEADERSHIP

University of South Florida

Tampa, FL

**Teaching Assistant @ Dept of Physics.**

04/2023 - Current

- **Mentored** 30+ first-year engineering students, **enhancing** their **understanding of Physics** through interactive weekly sessions.
- Conducted exam-prep sessions, resulting in an **average 20% grade improvement** for participating students.

# Jaemin Yang

*PhD Intern with Expertise in Data Science and Risk Analysis*

## Contact Information

Address: 158 Paddock Drive  
East, Savoy, IL, 61874  
Phone: +1-217-419-7821  
Email: jaeminy3@illinois.edu

## Skills

Python, R, PostgreSQL,  
NetLogo, Java, Machine  
Learning, Deep Learning,  
Statistical Analysis, Risk  
Analysis, Data Analytics

## Languages

Korean  
English

## Awards & Honors

- Best paper award from the journal of Nuclear Engineering and Technology, 2021
- Awarded for the best graduate (master's) student in research from engineering college, 2019
- Awarded for the best research achievement, Energy Research Center project, Korea Advanced Institute of Science and Technology (KAIST), 2017

## Professional Summary

PhD student in Informatics with a focus on machine learning and data analytics. Currently interning at Hyperlink, leading AWS and PostgreSQL projects. Over 7 years of experience in applying advanced data analytics in safety-critical domains. Known for innovative research integrating machine learning into risk assessment models.

## Work Experience

### Research & Teaching Assistant

Aug 2019 - Present

University of Illinois at Urbana-Champaign Urbana, IL

- Application of LLM to develop risk assessment model [**Python**]
- Developing a Breast Cancer Risk Assessment (BCRA) model from toxicant exposure using SEER, Comparative Toxicogenomics Database (CTD), and a variety of survey data (NHANES, IPUMS, ...) [**R, Python, Statistical Analysis**]
- Published the article associated with unknown accident identification using machine learning, and uncertainty analysis in the application of machine learning [**Machine Learning, Deep Learning, Python**]
- Developed a simulation using Netlogo for implementing fire operator movement [**NetLogo**]

### Full Stack Software Engineer Intern

May 2023 – Present

Hyperlink San Francisco, California

- Anomaly detection [**Python, PostgreSQL, Machine Learning**]
- Retention rate (%) analysis over weekly time periods [**Data Analytics, Statistical Analysis, PostgreSQL, Python**]
- Drawing conversion funnel [**Google Analytics, Data Analytics, Statistical Analysis, PostgreSQL, Python**]
- Documentation for AWS Database design [**GitHub, PostgreSQL**]

## Publications (Journal)

- **Jaemin Yang** et al., " Uncertainty analysis on support vector machine for measuring organizational factors in probabilistic risk assessment of nuclear power plants" Progress in Nuclear Energy. (2022): 104411
- **Jaemin Yang**, Jonghyun Kim. "Accident diagnosis algorithm with untrained accident identification during power-increasing operation." Reliability Engineering & System Safety (2020): 107032.
- Kim, Jonghyun, Deail Lee, **Jaemin Yang**, and Subong Lee. "Conceptual design of autonomous emergency operation system for nuclear power plants and its prototype." Nuclear Engineering and Technology 52, no. 2 (2020): 308- 322.
- **Jaemin Yang**, Jonghyun Kim. "An accident diagnosis algorithm using long short-term memory." Nuclear Engineering and Technology 50.4 (2018): 582-

588.

- **Jaemin Yang**, Taewan Kim, and Jonghyun Kim. "Analysis of errors of commission for a CE type plant with the advanced control room in the full power condition." Annals of Nuclear Energy 105 (2017): 184-195.

#### Conferences (Oral Presentations only)

- **AAAI-24 conference (submitted)**
- Two presentations in ESREL2020 PSAM 15, Venice, Italy
- 2020 ANS winter meeting and Nuclear Technology Expo, Chicago
- AHFE 2019, Washington D.C
- AHFE 2018, Orlando
- ISOFIC 2017, Washington D.C
- ESREL 2017, Portoroz, Slovenia
- Seven presentations in domestic (Korean) conferences

#### Education

Ph.D., Informatics (Data Analytics & Information Visualization) Aug 2019 - Present  
University of Illinois at Urbana-Champaign Urbana, IL

Master, Nuclear Engineering Mar 2017 - Feb 2019  
Chosun University, South Korea

Bachelor, Nuclear Engineering Mar 2011 - Feb 2017  
Chosun University, South Korea

# Sean(Xiangxi) Tian

11 Lechase Dr, Apt D, Gates, NY, 14606 • (929)-256-9067 • xtian8@ur.rochester.edu

LinkedIn: <http://www.linkedin.com/in/sean-t-326854203>

GitHub:<https://github.com/seant2436>

---

## OBJECTIVE

*Current Data Science student, graduating December 2024, with hands-on experience in data analytics, data mining, and statistical modeling. Seeking an entry-level Data Scientist role to apply my academic knowledge and further develop my skills in a professional setting, contributing to data-driven solutions and insights.*

---

## EDUCATION

*University of Rochester, MS in Data Science, Expected Dec 2024*

---

*University at Albany, State University of New York, Bachelor of Science, Actuarial Science and Mathematical Science, Minor in Economics, May 2023*

*Honors: Dean's List 7 Semester, 2020 Spellman Academic Achievement Awards, 2022 Spellman Academic Achievement Awards, Yang Family Scholarship, Hannaford EOP Scholarship, Ben V Smith Science Scholarship, Vanderzee Scholarship*

---

## SKILLS

### Programming and Data Analysis:

Python Libraries: NumPy, Pandas, Matplotlib, Scikit-Learn, Jupyter

R Libraries: dplyr, ggplot2, caret

Big Data Technologies: Hadoop, Spark, Hive

Deep Learning Libraries: TensorFlow, PyTorch

Database Management: MySQL (Database design, query optimization, data manipulation)

SAS: Proficient in data analysis and statistical modeling

MapReduce: Familiar with MapReduce for distributed data processing

### Development Tools and Practices:

Bash: Command line scripting and automation

Linux: Operating system and server administration

Version Control: Git for collaborative software development

### Languages:

Chinese Mandarin: Native/Bilingual Proficiency

English: Native/Bilingual Proficiency

---

## PROJECT EXPERIENCE

## *Exploring Temperature Relationships with Linear Regression Models*

- Employs linear regression models to examine temperature relationships.
  - Predicts "Apparent Temperature (C)" based on "Temperature (C)" and displays a high R-squared value, though an outlier is identified in the dataset.
  - Forecasts "Temperature (C)" using "Wind Speed (km/h)" and "Pressure (millibars)" as predictors, with accompanying summaries and plots for evaluation.

## *Text Generation using Markov Chains*

- Generate comedic text for a chosen comedian based on their comedy routines
  - Markov chains are employed on the occurrence of previous events, making it suitable for text generation
  - Data handling, cleaning, preprocessing, and Markov model construction

VR Game Database Management

- Create a database system for a VR game company to efficiently manage user profiles
  - Build a robust database for handling growing game data and simplify data management and analysis.
  - Gain player insights for improved gaming experiences and monitor company performance through data auditing.

Geospatial Data and Deep Learning

- Addresses the challenges of integrating geospatial data into deep learning workflows
  - Automates the conversion of georeferenced images and annotations for machine learning applications.
  - By bridging the gap between geospatial data and deep learning, this initiative supports a wide range of geospatial projects.

## ACADEMIC EXPERIENCE

*Teaching Assistant, University at Albany, SUNY, Albany, NY*

January, 2020 - May, 2021

- Conducted Calculus I Teaching Assistant for undergraduate level-Calculus I
  - Introduced the topics of Limits and Derivatives, Differential Rules, Application of Differentiation
  - Held office hours for one-on-one tutoring via Zoom

*ARM Mentor, University at Albany, SUNY, Albany, NY*

February , 2021 - May, 2021

- Mentored students in academic need for specific individual experienced academic difficulties
  - Provided essential information to students who need collaboration with campus resources
  - Demonstrated positive leadership

# Jesse Na

19 Pickwick Drive, Marlton, NJ 08053

856.313.7100

najess33@students.rowan.edu

## RESEARCH EXPERIENCE

### **Hoy Computational Research Lab at Rowan University**

Graduate and Undergraduate Research Assistant

**Glassboro, NJ**

*September 2022 - Present*

- Analyzed methyl ketone to carboxylic acid derivative transformation data with experimental tautomerization products utilizing machine learning potentials.
- Examined DNA and RNA nucleobases, resin derivatives, and polyaromatics on solid state metal electrodes to determine electron interactions and transport correlations.
- Utilized Python to improve Rowan University's Transport Code (pyRUQT) and other computational quantum chemistry software such as: GaussView, TorchANI, VMD, and Maestro.

### **Krummenacher Viral Pathogenesis Lab at Rowan University**

Undergraduate Research Assistant

**Glassboro, NJ**

*May 2022 - September 2022*

- Investigated the inhibition of bioactive peptides on feline coronavirus and herpes simplex virus.
- Maintained and established Vero and Crandell-Rees Feline Kidney Cell cell lines to assess cytotoxicity using CellTiter Blue cell viability and plaque reduction assays.
- Dissected and separated venom reservoirs from the sting apparatus of *Apis mellifera*.

## WORK EXPERIENCE

### **Virtua Healthcare Systems**

Pharmacy Technician

**Marlton, NJ**

*August 2019- September 2021*

- Assisted the Director of Pharmacy with keeping the pharmacy in compliance with USP 800 guidelines and ensuring proper removal of expired medications.
- Compounded IV medications and trained incoming pharmacy technicians on aseptic techniques.
- Utilized EPIC to assist nurses in analyzing patient profiles to help prevent drug interactions.
- Made regularly scheduled rounds to deliver medications to Nursing Units and treatment areas, restocked Pyxis, and collected controlled substance requisitions.

### **Rite-Aid Pharmacy**

Pharmacy Technician

**West Deptford, NJ**

*June 2019- September 2021*

- Compounded patient specific medications by calculating strength and dosages, weighed ingredients on a torsional balance, and added flavors or infused appropriate diluent.
- Administered seasonal influenza vaccines to patients under the supervision of a pharmacist.
- Processed insurance claims, transferred prescriptions, and assisted patients with prescriptions.

## EDUCATION

### **Rowan University**

B.S. in Biochemistry

**Glassboro, NJ**

*May 2023*

M.S. in Pharmaceutical Sciences

*expected December 2024*

### **Rowan College at Burlington County**

A.S. in Computer Science

**Mount Laurel, NJ**

*expected December 2024*

## LICENSES & CERTIFICATIONS

Google Advanced Data Analytics Professional Certification

*2023-Present*

IBM Machine Learning Professional Certification

*2023-Present*

Registered Pharmacy Technician

*2019-Present*

Certified Pharmacy Technician

*2019-2021*

# Benjamin Zhao

5 Van Marter Ct. Princeton, NJ 08540

Email: benz3927@gmail.com

<https://github.com/benz3927>

---

I am a rising sophomore undergraduate with project experience in data ingestion, coding algorithms, and statistical analysis. I am passionate about data science and eagerly pursuing an internship as a data analyst.

## EDUCATION

**Hamilton College**, Clinton, NY

Aug 2022 – Present

Cumulative GPA: 3.8

Key Coursework:

- 2023-2024: Linear Algebra, Probability and Statistical Inference, and Economic Theory and Evidence
- 2022-2023: Statistical Modeling and Applications, Multivariable Calculus, CPSCI-102 Design Principles

## TECHNOLOGIES

- Python, R, C++, UNIX, SQL, Microsoft Office
- DBeaver, AWS Athena, AWS EC2, Cloud Architecture

## PROJECTS

**Entity Match for Healthcare Providers** (GitHub: <https://bit.ly/3QcY87O>)

August 2023

- Ingested NPPES and PECOS data into AWS Athena tables.
- Used DBeaver to access Athena tables and ran SQL queries to clean data and conduct EDA.
- Wrote algorithm in Python to identify provider matches without using NPIs: paired rows from PECOS and NPPES datasets and compared key combinations of identifiers including dob, last name, address line 1, phone number, and first name to identify possible candidates.
- Conducted error analysis on unmatched pairs, explored different key combinations, and improved matching accuracy.
- Calculated weighted similarity score for pairs using Damerau-Levenshtein distance to filter matches.
- Correctly found 77.73% of the 3328 possible provider matches.

**24 Game in C++** (<https://github.com/benz3927/24>)

May 2023

- Created 24 Game in solo free play and multiplayer competition mode in C++.
- Utilized recursion to generate all possible combinations of arithmetic operations on pairs of numbers to solve 24 given 4 random or user-selected numbers.
- Built pointer-based stack classes for characters and doubles to handle parentheses and remove redundant brackets in string math expressions.
- Selected to demo a live play with the class.

**Web Scraping NHS Disease Symptoms** (GitHub: <https://bit.ly/3Y6tBup>)

Jan 2023

- Used Python request and Beautiful Soup packages to scrape a full list of NHS Diseases from A to Z.
- Caught all symptom bullet points under disease headers.
- Compiled all diseases and corresponding symptoms into a csv file.

**Examining Feature Rank and Dependency in the Q-CHAT-10 ASD Questionnaire**

Dec 2022

- Applied Random Forest in R to rank feature importance in ASD screening survey dataset collected by Dr. Thabtah.
- Utilized bootstrapping to group questions by relative feature importance and conducted chi-square tests between high and low-ranked questions.
- Identified dependency among similarly ranked variables and no statistical relationship between high and low-importance groups.
- Suggested improving survey methodology by selecting content-independent questions for early ASD identification, expanding scientific research, and exploring alternative survey questions.
- Submitted a manuscript to USCLAP (Undergraduate Statistics Class Project) Competition.

## EXPERIENCES

<b>Tennis Coach</b> , Princeton, NJ	Jun - Jul 2023
- Mentored young tennis player in tennis fundamentals.	
<b>Produce Stocker</b> , Whole Earth Center, Princeton, NJ	Jul 2022 – Aug 2022
- Stocked produce, prepped vegetables, worked on the produce display, and behind the cashier register.	
<b>Pawprint Newspaper Writer, Editor, and Webmaster</b> , Montgomery High School	Sep 2018 – Jun 2022
- Edited articles for sports, entertainment, and perspectives pages.	
- Initiated and transitioned newspaper to e-publish. ( <a href="https://themontgomerypawprint.wordpress.com">https://themontgomerypawprint.wordpress.com</a> )	

## HONORS

- 2nd place winner of the USCLAP of the American Statistical Association	Mar 2023
- Hamilton College Dean's List	Sep 2022 - May 2023

## MANUSCRIPTS AND PUBLICATIONS

- Zhao, Benjamin, and Drake Gorecki. "Examining Feature Rank and Dependency in the Q-CHAT-10 ASD Questionnaire." <https://bit.ly/3Oa8EtX>. 2023.
- Zhao, Benjamin. "The Minimum Wage — A High School Student's Perspective, a Comprehensive Review of Recent and Classical Literature." Curieux Academic Journal. October 2021.
- Zhao, Benjamin, Stefan Elez, Aarush Gupta, Shrivatsa Thukur. "Superstar' Firms, SMEs, and Inequality during the COVID-19 Pandemic." Young Global Scientists Journal 5th Edition. July 2021.

ATHARSS PRASATH  
2340 Piedmont Ave, Berkeley, CA 94704 ♦ (650)-730-8676 ♦ atharss.prasath@berkeley.edu

## SUMMARY

Dedicated and results-driven computer science student with a strong academic foundation in data structures, algorithms, design patterns, and object-oriented programming. Proven track record of delivering innovative and impactful software engineering projects.

## EDUCATION

---

<b>University of California, Berkeley, CA</b>	<b>May 2026</b>
Bachelor of Science in Electrical Engineering and Computer Sciences (EECS)	GPA: 3.80/4.00

**Academic Honors:** Semester Honors, awarded to top 20% of college of engineering students

### Relevant Coursework:

- CS 61A – Struc. & Interpretation of Comp. Programs (Python)
- CS 61B – Data Structures & Algorithms (Java)
- CS 70 – Discrete Math and Probability Theory
- CS 188 – Artificial Intelligence
- MATH 54 – Linear Algebra & Differential Equations
- IND ENG 162 - Linear Programming

## SKILLS

---

<b>Programming Languages:</b> Python, Java	<b>Frameworks:</b> Spring, REST API	<b>Packages:</b> Numpy, Matplotlib, Pandas, Scipy
<b>Relational Databases:</b> SQL, Postgres	<b>Tools:</b> MS Office, Git, Docker, JIRA, Jenkins	<b>Concepts:</b> Agile, SDLC, TDD, DevOps, CI/CD

## WORK EXPERIENCE

---

<b>Stotle Inc.</b> <i>Software Engineering Intern, Platforms Team</i>	<b>May 2023 - Aug 2023</b>
<ul style="list-style-type: none"><li>• Led R&amp;D projects within fast-paced technology start-up; re-engineered Java back-end of proprietary platform to improve page load times by 20%; iteratively analyzed root causes, implemented fixes, and adopted test driven development principles</li><li>• Reduced engineering debt by refactoring code; transitioned key features to a no-code configuration architecture to enable greater customization leading to 10% improvement in customer NPS</li><li>• Analyzed and benchmarked platform capabilities against top 3 competitors; engaged with cross-function stakeholders, incorporated feedback, and presented recommendations to accelerate generative AI roadmap by 6 months</li><li>• Improved dev ops by streamlining Google Cloud (GCP) implementation processes, leading to 15% reduction in post release issues; adopted Agile processes in a CI/CD delivery environment leveraging Docker, GitHub, and Jenkins</li></ul>	

## HIGHLIGHTED PROJECTS

### Voice-Controlled Robot Car (*Hardware-Software Integration, Python*)

- Built voice-controlled mobile robot car and designed motion controls using Arduino; finetuned Python classification algorithm to improve accuracy of the voice processing module
- Voice-controlled car is capable of interpreting verbal commands, and switching to the appropriate driving mode

### Equity Asset Price Predictor (*Python, LSTM, Data Science*)

- Created 50-neuron Long Short-Term Memory (LSTM) stock price prediction model using historical data from Yahoo Finance
- Achieved 70% improvement in prediction accuracy through multi-layer LSTM architecture and Min-Max scaling
- Visualized predictions using Matplotlib, with flexibility to analyze 5+ what-if scenarios

### Dynamic World Exploration Engine (*Data Structures, Java*)

- Designed a 2-D tile-based world exploration game in Java, applying data structures and shortest path algorithms
- Developed a highly complex and rule-based random world generation algorithm, incorporating 25+ constraints; showcased problem-solving skills and creativity, resulting in a dynamic UI and immersive user experience

## LEADERSHIP & EXTRACURRICULAR ACTIVITIES

---

<b>Chaplain @ Cal Acacia Fraternity</b>	<b>Nov 2022 - Present</b>
<ul style="list-style-type: none"><li>• Recruited and educated 50+ prospective members, and organized 10+ fraternity programs and events</li><li>• Tutored 20+ freshman and sophomore members on lower division CS courses including Data Structures and Algorithms</li></ul>	
<b>Member, Professional Development Committee @ HKN EECS Honors Society</b>	<b>Oct 2023 - Present</b>
<ul style="list-style-type: none"><li>• Develop professional development resources and organize career readiness events for 200+ strong Berkeley chapter of Eta Kappa Nu (HKN), national Electrical and Computer Engineering honor society; members are top 25% of EECS majors at Berkeley</li></ul>	
<b>Participant @ Cal Hacks Hackathon</b>	<b>Oct 2023</b>
<ul style="list-style-type: none"><li>• Participate in largest collegiate hackathon; collaborating as part of a 2-member team to build an automated prompt scoring algorithm with an API endpoint that can be leveraged by any web application reliant on LLMs</li></ul>	

# Sai Akhilesh Veldi

Arlington, USA | [saiakhileshv@gmail.com](mailto:saiakhileshv@gmail.com) | Portfolio | [Linkedin](#) | [Github](#) |

## EDUCATION

---

### Velocity Junior College

Telangana State Board of Intermediate Education(TSBIE)  
Percentage: 90.4%

June 2016 - Mar. 2018

### KL University, Hyderabad

Bachelor of Technology, Computer Science Engineering  
CGPA: 8.31/10

Hyderabad, India

Courses: Database Management Systems, Big Data Analytics, Machine Learning, Artificial Intelligence, Operating Systems

2018-2022

### The University of Texas at Arlington

Masters, Computer Science Engineering  
Courses: Design Analysis and Algorithms, Data Analysis and Modeling Techniques, Software Engineering

Arlington, USA

2022-2024

## PROFESSIONAL EXPERIENCE

---

### RESEARCH ASSISTANT

- Currently working on a Research project Under Dr. Erick C. Jones Jr.
- Data Analysis through Power BI, Tableau, python.
- Developed and Deployed Models through ML and Data Science Techniques

Arlington,  
October 2022 -  
Current

### SPARKS FOUNDATION

#### DATASCIENCE AND BUSINESS ANALYTICS INTERN:

- Worked on a Real-Time Project
- Developed Supervised Machine Learning model through K-Nearest Neighbors Algorithm.
- Predicted Student marks based on no of Student Study Hours.

#### WEBSITE DESIGNING, DIGITAL MEDIA MARKETING

- As a Tech enthusiast, I have created websites for different companies which include diversified sectors from Finance to IT, I have mainly used WordPress in the front end.
- FINANCE
- I.T
- ARCHITECTURE

## ACADEMIC PROJECTS

---

#### [Effect of Covid-19 on Education sector\(students\) in India](#)

Feb - May 2021

- Collected the data from scratch through Google forms.
- Deployed Tableau for preprocessing the collected data.
- Visualized through Matplotlib and Seaborn modules in Python Idle.
- Analyzed the data and concluded the results from our findings.

#### [Forecasting and Binning of Air Quality Index values towards encouraging green vehicles for sustainability.](#)

July - Nov. 2021

- As a Team Leader, I divided and allocated the work into 3 parts.
- Preprocessed the data to remove noisy data.
- Performed EDA for identifying key attributes.
- Evaluated for Past years to examine the accuracy rate, then forecasted for the future.
- Binned into six buckets: Moderate, Satisfactory, Good, Poor, Very Poor, Severe.
- Suggested alternative as Electric Vehicles for traditional vehicles.

#### [Plagiarism detection in Online Exams.](#)

Feb - Apr. 2022

- For training and testing, I created an online exam through the Google Forms and collected mouse movements and keystrokes where few are positive dataset and a few negative datasets
- Then I used machine learning techniques of different classification and different clustering algorithms on the collected data,
- At last, I evaluated the result and if necessary redo the data collection over and over until I made final decisions and conclusions accordingly. I used data augmentation to address the problem of underfitting the model

## SKILLS

---

**Technical Skills:** C, C++, Java, Python, SQL, Tableau, Pega, WordPress, Microsoft Excel, Word, Microsoft suite, Matlab.

## CERTIFICATIONS

---

Artificial Intelligence Engineer course in Simplilearn.

09/2020-12/2020

03/2021-04/2021

Introduction to Big Data course at Coursera.

NASSCOM Skillup Certificate on foundations of Artificial Intelligence.

07/2021-09/2021

Pega Certified System Architect

02/2021-07/2021

## LEADERSHIP

---

### Head

Badminton wing, Sports club

- Responsible for hosting Annual sports fest.
- Nurtured a team of 8 members for Badminton Doubles.

## AWARDS & ACHIEVEMENTS

---

- Awarded 2<sup>nd</sup> prize in inter-university badminton tournament, 2019
- Ranked 5<sup>th</sup> in Hackathon held at our university, 2020

# Kourosh T. Baghaei

Fairfax, VA 22032 |  (618)303-2749 |  k.teimoori71@gmail.com |  linkedin.com/in/kourosh-t-baghaei |  
 [github.com/kbaghaei](https://github.com/kbaghaei) |  <http://baghaei.me/>

## Applied Research Scientist, Machine Learning Engineer

### TECHNICAL SKILLS

**Programming Languages:** Python, C/C++, SQL, Java, C#, R, Go, Swift, Bash

**Libraries:** PyTorch, Keras, TensorFlow, NumPy, NLTK (Natural Language Processing), Pandas, Scikit-learn, OpenCV

**DevOps Tools:** Git, Docker, Kubernetes, AWS

**Data Base Systems:** MySQL (Relational Database), PostgreSQL (Relational Database), Redis (NoSQL)

**Machine Learning:** Neural Networks, Natural Language Processing, Reinforcement Learning Specialization

### PROFESSIONAL EXPERIENCE

#### Graduate Research Assistant

01/2022 – Present

*George Mason NLP Group*

*Fairfax, VA*

- Development of conversational Vision and Language Navigation agent for urban environment (Python, PyTorch) by leveraging large language models (LLMs)
- Development of cloud-based multiplayer game for crowd-sourced annotation of urban navigational routes (Unity3D, C#, AWS )

#### Machine Learning Engineer

05/2020 – 12/2020

*National Strategic Planning and Analysis Research Center( NSPARC)*

*Starkville, MS*

- Development of data pipeline to capture human pose from 2D images (**Computer Vision**) on iOS mobile device.
- 3D human pose estimation (3D Imagery) from 2D pose information using PyTorch and Swift.

#### Graduate Research Assistant

09/2018 – 05/2020

*Mississippi State University*

*Starkville, MS*

- Development of data pipelines in order to gather, extract, and prepare **Time series data** of around 60,000 ICU stays stored on **PostgreSQL** (Big Data)
- Design and implementation of an **Interpretable** attention-based deep neural network in order to predict the outcome of ICU stays using **PyTorch** framework.
- presented and published results in an [IEEE conference on Artificial Intelligence and Machine Learning](#).

### PROJECTS

#### Interpretable prediction of the outcome of ICU stays, by a deep recurrent network | *PostgreSQL, PyTorch*

- Development of an attention-based GRU model using **PyTorch** to make predictions about the outcome of ICU stays. Based on the time series data of the ICU stays in hospital. (Master's thesis project).

#### Controlling the character with 2D Poses extracted from webcam | *OpenCV, PyTorch, Unity3D, C#*

- Development of a 2D video game in Unity, that extracts the body pose of the player from the webcam's video stream using **pose estimation** and use that info to control the character in the game.  
([click here to watch the demo](#))

#### Sentiment Analysis of Movie Reviews | *NLTK, sklearn, Pandas*

- Data preprocessing, cleaning, and feature reduction of comments of movie reviews.
- Development of predictive models to perform **sentiment classification** of natural language texts.

#### A minimal Implementation of Bert | *PyTorch, Python*

- Development of the Self-Attention model and the Bert architecture in PyTorch

### EDUCATION

#### George Mason University

Fairfax, VA

*PhD in Computer Science*

*01/2021 – 12/2025*

#### Mississippi State University

Starkville, MS

*Master of Science in Computer Science (GPA: 4.0 / 4.0)*

*08/2018 – 12/2020*

#### K.N. Toosi University of Technology

Tehran, Iran

*Bachelor of Science in Software Engineering*

*08/2010 – 08/2015*

**Yousuf Qari**  
Somerset, NJ  
[ybqari@gmail.com](mailto:ybqari@gmail.com)  
(908) 307-9084  
[linkedin.com/in/yousuf-qari/](https://linkedin.com/in/yousuf-qari/)

---

## **EDUCATION**

**Temple University**, Philadelphia, PA

*Bachelor of Science*

*Majors: Computer Science, Neuroscience*

Expected Graduation: May 2025

GPA: 3.89/4.0

## **PROFESSIONAL EXPERIENCE**

**Genellipse Inc.**, *Data Engineer Intern*

May 2023-Present

- Revamped an ETL workflow using Python, Pandas library and Azure DevOps, achieving a 100% success rate across the workflow, an overall increase of 34% efficiency in data processing functions.
- Leveraged Microsoft Azure services to monitor failures in 11 functions to identify and resolve issues, and employed data-driven insights to implement targeted improvements, saving the company from upgrading to a more expensive Azure subscription plan
- Designed and developed four report pages using Power Query Editor and DAX formulas in Power BI containing information-rich bar charts and tables, providing actionable business intelligence used and distributed by the client's Treasurer's Office
- Met weekly with client and company leadership to present team's progress on production support and change requests for client-requested reports

**Temple University Department of Computer Science**, *Research Assistant*

May 2022-Present

- Researched with Professor Xinghua Mindy Shi on machine learning prediction models with MRI brain data from patients with Parkinson's disease at Temple University hospital
- Contributed to data cleaning for machine learning algorithms which use patients' genomic data to train a model to predict the stage of cancer for prostate cancer patients using Python and ML libraries such as Pandas and NumPy

**Temple Science & Education Research Center**, *Computer Lab Consultant*

August 2022-Present

- Consulted 50+ users with technical problems concerning a wide variety of issues including Microsoft Office, printing & scanning, media editing, web publishing, and hardware configuration
- Closely monitored & maintained order in the lab for 10+ hours a week by enforcing lab policies, ensuring general computer function, stocking supplies, & proactively addressing unexpected issues

## **PROJECT EXPERIENCE**

**Temple University OwlHacks 2023 Hackathon - [Skribble](#)**

- Designed and developed a cross-platform mobile app with two teammates using React Native on Expo that serves as a text-based mental-health journaling app geared towards children aged 6-12, with the goal of establishing self-care practices through a child-friendly user interface
- Received the first-place award in the Health and Wellness category and received the Hackathon-wide award for Best User Experience Design among 150+ participants
- Submitted a presentation to event judges and industry representatives that outlines the mission, demos app functionality and summarizes the development/design process

**Mythical Creatures Game**

- Designed and developed a 2-player mythical card game in Python which deals each player two cards from a generated deck of 20, each with unique attack and defense points, and gives both players an option to trade one of their cards before deciding a winner based on the cards' total scores.

## **ADDITIONAL INFORMATION**

**Languages:** Python, Java, React Native, SQL

**Skills:** Microsoft Azure, Power BI (data visualization, Power Query Editor, DAX), Git, Pandas library

# Preston DeLeo

pcd42@case.edu • (203) 586-8538 • linkedin.com/in/preston-deleo • github.com/prestondeleo

## EDUCATION

---

**Case Western Reserve University | Cleveland, OH | Expected Graduation:** May, 2025

**Majors:** Computer Science and Cognitive Science | **Minors:** Artificial Intelligence and Mathematics

**Honors:** Dean's Honors List

**Relevant Courses:** Data Structures, Algorithms, Artificial Intelligence, Machine Learning, Probability, Linear Algebra

## SKILLS

---

**Programming Languages:** Java, Python, Julia, C#, SQL, R, JavaScript, HTML, CSS

**Libraries/Frameworks:** Pytorch, Tensorflow, Scikit-learn, Matplotlib, Numpy, JUnit, PyTest, Node.js

**Tools:** Git, Bitbucket, GitHub, Jira, Linux, HPC, VS Code, Jupyter Notebooks, Microsoft Office

**Software Development:** Agile methodology, Object-oriented design, Version control

## PROFESSIONAL EXPERIENCE

---

**Undergraduate Researcher | SDLE Research Center | Cleveland, OH**

**January 2023 - Present**

- Developing a deep learning model to categorize and predict 6000+ crystallite systems with over 30 features
- Co-authoring multiple research papers and on track to publish six papers by the end of undergraduate studies
- Performing data visualization, machine learning, code automation, debugging, and building software packages

**Data Analyst Intern | CREF | Boston, MA**

**June 2022 - July 2022**

- Conducted data cleaning on hundreds of CSVs, Excel spreadsheets, and unstructured text data files to address missing values, outliers, and inconsistent data types
- Applied data transformation on 100+ files including normalization, scaling, aggregation, and string manipulation
- Crafted SQL queries for metadata in databases, creating data dictionaries of 1000+ elements and 20+ features

## PROJECTS

---

**Crystallite Growth Analysis | Research Project**

**June 2023 - Present**

- Pioneering the first method of analysis and software package for crystallite growth properties in Python
- Building a spatiotemporal graph neural network to predict growth patterns of 6000+ crystallites on 500+ timesteps
- Carried out tasks such as pre-processing data, performance analysis on the model and 10+ clustering algorithms

**8-Puzzle Game | Personal Project**

**September 2022 - October 2022**

- Implemented an 8-puzzle game along with an AI solver in Java, adhering to object-oriented programming conventions and executed proper testing methodologies and analysis of performance
- The AI solves any 8-puzzle utilizing A\* and local beam search algorithms to find the shortest path to the solved state
- Handles 10,000+ puzzle states at once and determines unsolvable states by calculating the parity of a permutation

## LEADERSHIP EXPERIENCE

---

**Vice President | Case Western Association for Computing Machinery**

**August 2022 - Present**

- Chaired executive meetings, delegated tasks, and restructured organization so executives have an equal say in affairs
- Orchestrated campus hackathons, including an innovative quantum computing hackathon that introduced hundreds of students to cutting-edge concepts in an emerging field
- Managed programming competitions, giving students a platform to apply skills with prizes that totaled over \$2000
- Coordinated 10+ engaging talks featuring guest speakers in the tech industry for students
- Delivered 5 presentations on various topics in tech, such as a comprehensive overview of neural networks

# **Yi Zhang**

Tucson, Arizona 85705

[yizh2018@gmail.com](mailto:yizh2018@gmail.com) | (213)-245-9556 | [Publication](#) | [LinkedIn](#)

## **EDUCATION**

08/2021 - 05/2025

**University of Arizona**

Tucson, AZ

*Ph.D. in Biostatistics GPA: 4.0/4.0*

Relevant Courses: Statistical Machine Learning, Neural Networks, Big Data Technologies, Longitudinal and Mixed Effects Models, Theory of Statistics, Theory of Linear Models, Statistical Consulting

08/2017 - 05/2019

**University of Southern California**

Los Angeles, CA

*M.S in Biostatistics GPA: 3.82/4.0*

Relevant Courses: Advanced Statistical Computing, Experimental Designs, Machine Learning for Health Sciences

09/2011 - 07/2016

**Chongqing Medical University**

Chongqing, China

*Bachelor of Medicine in Clinical Medicine*

## **SKILLS**

R, Python, SQL, High-Performance Computing in Linux Environment, Git, Jupyter, Pandas, NumPy, Pandas, Scikit-learn, Keras, TensorFlow, Matplotlib, PySpark, SAS, SPSS, STATA

## **WORK EXPERIENCE**

01/2022 - Present

**Graduate Assistant, Research Part-time**

Tucson, AZ

University of Arizona

### **Unsupervised Learning Modeling on Time Series Data**

- Pioneered a new system describing the severity of sleep apnea by hierarchical clustering algorithm and dynamic time warping (DTW) distances, improving the identification of differences in survival;
- Optimized Cox regression model by LASSO and Elastic-Net for feature selection and area under the curve (AUC) for evaluation of model performance;
- Resolved high collinearity in the model by variance inflation factor (VIF) and evaluated model assumptions;

### **Functional Analysis on scRNA-seq Data**

- Removed batch effect in integrated scRNA data and inferred pseudotemporal trajectory of cells using minimum spanning tree approach after dimension reduction of gene expression by PCA and UMAP;
- Applied Generalized Additive Model (GAM) to study the association between diseases and estimated smoothed expression curves over pseudotime;

### **Statistical Analysis and Presentation**

- Upgraded R-Shiny App to develop reports and dashboard for COVID-19 statistics, enabling decision-making;
- Expedited the development of a new statistical method Zero Inflated Smoothing Splines (ZISS) model to improve performance of zero-inflated responses in high sparsity data;
- Collaborated with researchers in a cross-functional team to generate and present comprehensive statistical reports for non-statistical audiences;

09/2019 - 06/2021

**Statistician I Full-time**

Los Angeles, CA

University of Southern California

### **Modeling and Hypothesis Testing**

- Improved the performance of linear mixed model by spaghetti plots exploring factors related to smoking urges;
- Developed Autoregressive Integrated Moving Average (ARIMA) model to forecast the substance use post COVID-19, which was applied to investigate change of consuming behavior pre and post COVID-19;
- Performed hypothesis testing to assess the effect of different smoking behaviors on clinical outcomes;

### **Collaboration and Presentation**

- Provided professional consultations on statistical methodologies, data management, data collection, and reproducible programming to researchers with various backgrounds, facilitating 4 publications;
- Designed and led tutorials of feature engineering and model training for 20 junior analysts, the recordings of which are continually utilized as new staff training;

05/2018 - 08/2018

**Data Analyst Intern**

Shanghai, China

Hellobike

- Developed customized SQL query based on over 100 million records of personal and GPS information, improving 5% of efficiency in locating and recycling bicycles;
- Designed and implemented A/B testing to assess the impact of marketing strategies, resulting in a 6% higher weekly retention rate;

## **PROJECT EXPERIENCE**

10/2018 – 05/2019

**Supervised Learning Modeling on Satellite Image Data**

Los Angeles, CA

Graduate Thesis at University of Southern California

- Examined and classified land cover changes by Random Forest algorithm with 30 m resolution satellite imagery obtained from the Landsat Thematic Mapper;
- Calculated, analyzed and visualized the image changes in metrics that describing greenness;

# Wenjin Zhang

732-268-0001 • [wz315@soe.rutgers.edu](mailto:wz315@soe.rutgers.edu) • 11 Canterbury Circle, Somerset, NJ • 08873

---

## EDUCATION

### **Doctor of Philosophy-PhD Computer Engineering**

Rutgers University GPA: 4.0  
Major: Computer Engineering Expected Graduate: December, 2025

### **Master of Science**

Monmouth University GPA: 3.87  
Major: Software Engineering Graduate: December, 2019

### **Bachelor of Engineering**

Changshu Institute of Technology GPA: 3.85  
Major: Software Engineering Graduate: June, 2018

## PUBLICATIONS

- W. Zhang, X. Zang, L. Huang, Y. Sui, J. Yu, Y. Chen and B. Yuan. "DynGMP: Graph Neural Network-based Motion Planning in Unpredictable Dynamic Environments." 2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2023)
- C. Shi, T. Zhao, W. Zhang, T. Mahdad, Z. Ye, Y. Wang, N. Saxena, Y. Chen, "Defending against Thru-barrier Stealthy Voice Attacks via Cross-Domain Sensing on Phoneme Sounds", 42nd IEEE International Conference on Distributed Computing Systems (ICDCS), Bologna, Italy, 2022.
- W. Zhang, J. Wang and F. Lan, "Dynamic hand gesture recognition based on short-term sampling neural networks," in IEEE/CAA Journal of Automatica Sinica, vol. 8, no. 1, pp. 110-120, January 2021, doi: 10.1109/JAS.2020.1003465.
- W. Zhang, J. Yan and C. Yu, "Smart Parking System Based on Convolutional Neural Network Models," 2019 6th International Conference on Information Science and Control Engineering (ICISCE), Shanghai, China, 2019, pp. 561-566, doi: 10.1109/ICISCE48695.2019.00118.
- C. Yu, A. Lu, S. Cassidy and W. Zhang, "ActionTime: Help Children Develop Execution Skills", The 5th IEEE International Conference on Collaboration and Internet Computing, Los Angeles, California, USA, 2019
- W. Zhang and J. Wang, "Dynamic Hand Gesture Recognition Based on 3D Convolutional Neural Network Models," 2019 IEEE 16th International Conference on Networking, Sensing and Control (ICNSC), Banff, AB, Canada, 2019, pp. 224-229.

## WORK EXPERIENCE

**Company:** Rutgers University

**Position:** Research Assistant

**Location:** 96 Frelinghuysen Rd, Piscataway, NJ 08854

**Time Period:** Jan 2020 – Present

**Company:** Monmouth University

**Position:** Adjunct Professor

**Location:** West Long Branch, New Jersey, US

**Time Period:** April 2020 – December 2020

**Teaching:** CS/SE 698 deep learning with python (3 credit); CS 533 50 Database System Implement (3 credit).

**Company:** Vydia

**Position:** Software Engineer Intern

**Location:** New Jersey, US

**Time Period:** June 2019 - September 2019

**Description:** Refactoring the Docker file by replacing the Ubuntu Linux with Alpine Linux; Web backend development.

**Company:** Monmouth University

**Position:** Summer Research Mentor

**Location:** New Jersey, US.

**Time Period:** June 2019 – September 2019

## PROJECT DEMO

### **Dynamic Motion Planning**

#### **• Project Description:**

we developed DynGMP, a graph neural network (GNN)-based planner that provides high-performance planning

solutions in unpredictable dynamic environment

- **Paper** was accepted by IROS 2023

### **Hand Gesture Recognition**

- **Project description:**

This project is aimed to recognize the dynamic hand gestures using a webcam. The hand gesture types include swiping left, swiping right, swiping down, swiping up, pushing hand away, pushing hand in, turning hand clockwise, turn hand counterclockwise, thumb up, thumb down, shaking hand, drumming fingers, stop sign and so on.

- **Technique:** Python, PyTorch, OpenCV, PIL, CUDA, 3D Convolutional Neural Network, Flask, HTML/CSS, JavaScript/jQuery
- **Training dataset:** 20BN-jester Dataset V1
- **Paper:** <https://ieeexplore.ieee.org/document/8743159>
- **Project demo video:** <https://youtu.be/gBE7cOssUbI>, <https://youtu.be/bSkLH-Ng0D8>

### **Smart Parking System**

- **Project description:**

This project is aim to detect the empty parking space based on cameras.

- **Technique:** Python, Tensorflow, Keras, OpenCV, PIL, CUDA, Convolutional Neural Network, Flask, HTML/CSS, JavaScript/jQuery
- **Paper:** <https://ieeexplore.ieee.org/document/9107707>
- **Project demo video:** <https://youtu.be/wBjpJW9iG0Q>

Jesse Graham

[Email](#) | [LinkedIn](#) | 848-404-8570 | 1 Amherst Place Parlin, NJ 08859

**OBJECTIVE:** Currently seeking Internships and other opportunities to further my career in Computer Science and Cyber Security.

## **EDUCATION & CERTIFICATIONS**

## Rutgers University, New Brunswick, New Jersey

## Bachelor of Computer Science

## Minor in Data Science

GPA: 3.537/4.0      Expected graduation date: May 2026

**Google Cybersecurity Certificate, Coursera, 2023**

## **TECHNICAL & NON-TECHNICAL SKILLS**

- Java
  - Python
  - C++
  - Linux
  - SQL
  - SIEM Tools
  - Security Frameworks and Controls
  - Security Hardening
  - Microsoft Word, Excel, PowerPoint

## Projects

## One Shot Learning (Computer Architecture)

Fall 2023

- Used file I/O, dynamic memory allocation, and matrix manipulation to develop a C program that implements a simple machine-learning algorithm to predict house prices based on specific attributes using training data.

## EXPERIENCE

## Olive Garden, Piscataway NJ;

## Server

April 2023 – Present

- Provide exceptional customer service by effectively and efficiently assisting patrons with their orders, ensuring a positive dining experience.
  - Collaborate with kitchen and bar staff to expedite orders, resulting in reduced wait times and improved customer satisfaction scores.
  - Act as a point of contact for customer inquiries, resolving issues promptly and professionally, leading to a 10% increase in customer satisfaction ratings.

## **Aquatics and Fitness Center, Parlin NJ;**

## **Lifeguard & Swim Instructor      July 2019 – Present**

- Vigilantly ensure the safety of all facility attendees by maintaining constant surveillance and responding swiftly to any emergency situations.
  - Administer life-saving techniques and first aid as required, effectively mitigating potential hazards and ensuring the well-being of patrons.
  - Develop and implement individualized swim programs, resulting in a significant improvement in student skill development and satisfaction.
  - Collaborate with parents to provide progress updates, fostering positive relationships and creating a supportive learning environment.

## ACTIVITIES & AWARDS

- |   |                                                       |           |                          |
|---|-------------------------------------------------------|-----------|--------------------------|
| • | MEET (Minority Engineering Educational Task)          | Member    | September 2022 – Present |
| • | Undergraduate Student Alliance of Computer Scientists | Member    | September 2022 – Present |
| • | MetLife Foundation Pathways Scholarship Program       | Recipient | May 2023 — Present       |

## **RELEVANT COURSES**

- Intro To Computer Science
  - Data Structures
  - Calculus II
  - Intro To Discrete Structures
  - Computer Architecture
  - Basic Statistics for Research

## HIRA ANEES AWAN

Iowa City, IA 52246

(984) 238-6727 [hiraaneesawan@uiowa.edu](mailto:hiraaneesawan@uiowa.edu)

<https://hiraaneesawan.github.io/>

[www.linkedin.com/in/hiraaneesawan](http://www.linkedin.com/in/hiraaneesawan)

## PROFILE

A motivated and determined **Fulbright Scholar** and **Biostatistician** with substantial experience in multidisciplinary research related to immunology, cancer systems biology, and data science. Currently, a **Carver Fellow** and **PhD candidate** exploring the world of biomedical imaging and deep learning. Published author, recognized for leadership and problem-solving skills.

**Technical Skills:** R, Python, Bash, SAS (basic), MATLAB, C#, C, Java, JavaScript, Git, Docker

## EDUCATION

**THE UNIVERSITY OF IOWA**, College of Engineering, Iowa city, IA

**PhD in Biomedical Engineering**, May 2025. GPA: 3.93 Carver Fellowship

**Advisor:** Joseph Reinhardt

**DUKE UNIVERSITY**, School of Medicine, Durham, NC

**Master of Biostatistics**, May 2021. GPA: 3.968 Fulbright Scholarship - Partial Tuition Scholarship

**Master's project (Awarded best project award – Data Science category)**

A deep learning approach for joint batch calibration and clustering of multi-center flow cytometry data

**Advisors:** Cliburn Chan, Georgia Tomaras, Jichun Xie

**UNIVERSITY OF ENGINEERING AND TECHNOLOGY**, Lahore, PK

**Bachelor of Science, Major: Electrical Engineering, Minor: Computer Engineering**, May 2017. GPA: 3.7/4.0.

**Bachelor's project:**

Theatre for In Silico Oncology (TISON) – A Multiscale Cancer Systems Modeling and Simulation Platform

**Advisors:** Safee Ullah Chaudhary, Khalid Mahmood Hasan

## PROFESSIONAL EXPERIENCE

**THE UNIVERSITY OF IOWA**, Iowa city, IA

2021-Present

**The Reinhardt Biomedical Imaging Laboratory**

**Graduate Research Assistant**

- Working on a data analysis pipeline to associate biomechanical features of the lungs with disease outcomes
- Working on a deep learning architecture to predict severe exacerbations from computed tomography (CT) images in chronic obstructive pulmonary disease (COPD) patients
- Working on a statistical analysis pipeline to classify symptomatic and asymptomatic tobacco-exposed people with preserved spirometry

**THE DUKE HUMAN VACCINE INSTITUTE**, Durham, NC

2020-2021

**Tomaras Laboratory**

**Research Intern**

- Worked on a statistical analysis pipeline to unveil the effects of preexisting CMV immunity on vaccine-specific cellular and humoral responses and the risk of HIV acquisition in HVTN 505 Phase IIb Human Vaccine Efficacy Trial

**UNITED WE REACH**, Lahore, PK

2018-2019

**Curriculum Developer**

- Collaborated with coordinators of more than 500 primary schools across Pakistan to enhance the current curriculum of science and mathematics
- Transformed the computer science curriculum by adding modern programming languages for children like Scratch. For higher grades, incorporated basic Python coding exercises

**LAHORE UNIVERSITY OF MANAGEMENT SCIENCES**, Lahore, PK

2016-2018

**Biomedical Informatics Research Lab**

**Team Lead (2018)**

- Initiated Phase II of Theatre for in silico Oncology (TISON) – a multiscale cancer systems modeling and simulation web platform
- Worked with biologists towards incorporating more features to enhance the overall desirability of the software platform
- Developed and revamped three out of seven scales (Cell circuit design, Phenotype design, Tissue design) of the platform
- Implemented the simulation engine of the platform by integrating data from all the scales
- Trained a new team of four fresh undergraduate students and accelerated the development process by dividing scales among them

**Research Associate (2017)**

- Contributed towards the development of MATLAB-based Attractor Landscape Analysis Toolbox for Cell Fate Discovery and Reprogramming (ATLANTIS) and later enhanced the performance of the toolbox
- Developed the manuscript for the toolbox with my team and independently worked on supplementary material and user manual for the toolbox and later executed the rebuttal process with my team and dealt with a major revision within 30 days

**Research Intern (2016)**

- Gained proficiency in software development over the summer and familiarized myself with the latest trends and practices in the world of web development
- Initiated Phase I of Theatre for in silico Oncology (TISON) – a multiscale cancer systems modeling and simulation web platform and worked on cell circuit design scale to familiarize myself with the underlying software practices and later broadening my knowledge about the biology of cell cycles

**UNIVERSITY OF ENGINEERING AND TECHNOLOGY, Lahore, PK**

2015-2016

**Research Intern (2016)**

- Worked on a Linux based platform (Ubuntu) towards the development of Next Generation Sequencing (NGS) data analysis pipeline

**TEACHING ASSISTANT EXPERIENCE****DUKE UNIVERSITY, Durham, NC**

Fall 2020

- Course: Applied Biostatistical Methods I

**UNIVERSITY OF ENGINEERING AND TECHNOLOGY, Lahore, PK**

Fall 2015

- Course: Applied Biology

**PUBLICATIONS****JOURNALS**

W. McKleroy, T. Shing, W. H. Anderson, M. Arjomandi, **H. A. Awan**, et al., “Longitudinal Follow-Up of Participants With Tobacco Exposure and Preserved Spirometry.” *JAMA*. 2023; 330(5):442-453.

O. S. Shah, M. F. A. Chaudhary, **H. A. Awan**, F. Fatima, Z. Arshad, B. Amina, M. Ahmed, et al., “ATLANTIS – Attractor Landscape Analysis Toolbox for Cell Fate Discovery and Reprogramming.” *Nature Scientific Reports* 8, no. 1 (2018): 3554.

M. N. Gondal, A. Rehman, A. Arif, M. U. Sultan, **H. A. Awan**, Z. Arshad, M. F. A. Chaudhary et al., “TISON: A Next-generation Multi-scale Modeling Theatre for in silico Systems Oncology”, *Molecular Systems Biology*. (*In Preparation*)

**CONFERENCES & POSTER PRESENTATIONS**

**H. A. Awan**, M. F. A. Chaudhary, S. E. Gerard, A. P. Comellas et al., “Deep residual convolutional network predicts future severe exacerbations of COPD in SPIROMICS”, American Thoracic Society (ATS) Conference, May 2023, Washington DC, USA.

**H. A. Awan**, M. F. A. Chaudhary, W. Labaki, J. Guo, et al., “Novel QCT textures are associated with sRAGE and IL-6 in SPIROMICS”, European Respiratory Society (ERS) International Congress, September 2022, Barcelona, Spain.

**H. A. Awan**, C. J. L. Angel, G. Tomaras, “Investigating the impact of preexisting CMV immunity on the risk of HIV acquisition in HVTN505 vaccine efficacy trial for HIV”, *Duke Surgery 5th Annual Research Day*, April. 2021, Durham, NC, USA. (*abstract selected for oral presentation*)

**H. A. Awan**, C. J. L. Angel, J. McCarthy, G. Tomaras, “Preexisting CMV immunity is associated with blunted vaccine specific cellular and humoral responses and increased risk of HIV acquisition in phase IIb human vaccine efficacy trial”, *16th Annual Duke Center for AIDS Research Virtual Fall Scientific Retreat*, Sept. 2020, Durham, NC, USA. (*abstract selected for oral presentation*)

S. U. Chaudhary, M. N. Gondal, **H. A. Awan**, A. Rehman, A. Arif, R. Hussain, H. Khawar, Z. Arshad, M. F. A. Chaudhary et al., “A Next Generation Multi-scale Modeling Theatre for In Silico Oncology”, *14th International Conference on Systems Biology (ICSB 2020)*, Jan. 2020, Zurich, Switzerland.

M. F. A. Chaudhary, **H. A. Awan**, H. Hameed, Z. Arshad, M. Abubakar, K. M. Hasan, B. Wajid, et al., “TISON - Theatre for in silico Oncology: A Next-Generation Multiscale Modeling Platform for Predicting Cancer Growth, Development and Metastasis,” *DICE Mega Innovation and Entrepreneurship Summit*, Dec. 2016, National University of Science and Technology, Pakistan (*won second prize among 350 posters presented*)

M. F. A. Chaudhary, **H. A. Awan**, H. Hameed, Z. Arshad, M. Abubakar, K. M. Hasan, B. Wajid, et al., “TISON - Theatre for in silico Oncology: A Next-Generation Multiscale Modeling Platform for Predicting Cancer Growth, Development and Metastasis,” *BioSymposium 2016*, Lahore University of Management Sciences, Lahore, Pakistan (*honorable mention at the poster presentation contest*)

E. Serpedin, Z. Arshad, H. Hameed, M. F. A. Chaudhary, **H. A. Awan**, B. Wajid, “Investigating Functional Relationships of Five Critical Genes Present in the Familial Non-Synonymous Loci with Risk of having Acute Lymphoblastic Leukemia,” *Integrative Biology Conference*, July 18-20, 2016, Berlin, Germany

**HONORS & AWARDS**

- Awarded the prestigious ***Carver Fellowship*** to pursue PhD in Biomedical Engineering at The University of Iowa
- Awarded best master's project award (data science category) by the Department of Biostatistics and Bioinformatics at Duke University
- Awarded the prestigious ***Fulbright Scholarship*** to study Biostatistics at Duke University
- Granted 2nd prize for TISON - Theatre for in silico Oncology among 350 other projects at the DICE Mega Innovation Summit
- Awarded ***Student Research Grant*** for the project, "Diagnostic Tests for Cancer (now TISON)", by the National Grassroots ICT Research Initiative
- Awarded Dean's Honor Roll for Fall 2014, Spring 2015, Fall 2016, Spring 2017, Department of Electrical Engineering, University of Engineering and Technology, Lahore, Pakistan

## HIRA ANEES AWAN

Iowa City, IA 52246

(984) 238-6727 [hiraaneesawan@uiowa.edu](mailto:hiraaneesawan@uiowa.edu)

<https://hiraaneesawan.github.io/>

[www.linkedin.com/in/hiraaneesawan](http://www.linkedin.com/in/hiraaneesawan)

## PROFILE

A motivated and determined **Fulbright Scholar** and **Biostatistician** with substantial experience in multidisciplinary research related to immunology, cancer systems biology, and data science. Currently, a **Carver Fellow** and **PhD candidate** exploring the world of biomedical imaging and deep learning. Published author, recognized for leadership and problem-solving skills.

**Technical Skills:** R, Python, Bash, SAS (basic), MATLAB, C#, C, Java, JavaScript, Git, Docker

## EDUCATION

**THE UNIVERSITY OF IOWA**, College of Engineering, Iowa city, IA

**PhD in Biomedical Engineering**, May 2025. GPA: 3.93 Carver Fellowship

**Advisor:** Joseph Reinhardt

**DUKE UNIVERSITY**, School of Medicine, Durham, NC

**Master of Biostatistics**, May 2021. GPA: 3.968 Fulbright Scholarship - Partial Tuition Scholarship

**Master's project (Awarded best project award – Data Science category)**

A deep learning approach for joint batch calibration and clustering of multi-center flow cytometry data

**Advisors:** Cliburn Chan, Georgia Tomaras, Jichun Xie

**UNIVERSITY OF ENGINEERING AND TECHNOLOGY**, Lahore, PK

**Bachelor of Science, Major: Electrical Engineering, Minor: Computer Engineering**, May 2017. GPA: 3.7/4.0.

**Bachelor's project:**

Theatre for In Silico Oncology (TISON) – A Multiscale Cancer Systems Modeling and Simulation Platform

**Advisors:** Safee Ullah Chaudhary, Khalid Mahmood Hasan

## PROFESSIONAL EXPERIENCE

**THE UNIVERSITY OF IOWA**, Iowa city, IA

2021-Present

**The Reinhardt Biomedical Imaging Laboratory**

**Graduate Research Assistant**

- Working on a data analysis pipeline to associate biomechanical features of the lungs with disease outcomes
- Working on a deep learning architecture to predict severe exacerbations from computed tomography (CT) images in chronic obstructive pulmonary disease (COPD) patients
- Working on a statistical analysis pipeline to classify symptomatic and asymptomatic tobacco-exposed people with preserved spirometry

**THE DUKE HUMAN VACCINE INSTITUTE**, Durham, NC

2020-2021

**Tomaras Laboratory**

**Research Intern**

- Worked on a statistical analysis pipeline to unveil the effects of preexisting CMV immunity on vaccine-specific cellular and humoral responses and the risk of HIV acquisition in HVTN 505 Phase IIb Human Vaccine Efficacy Trial

**UNITED WE REACH**, Lahore, PK

2018-2019

**Curriculum Developer**

- Collaborated with coordinators of more than 500 primary schools across Pakistan to enhance the current curriculum of science and mathematics
- Transformed the computer science curriculum by adding modern programming languages for children like Scratch. For higher grades, incorporated basic Python coding exercises

**LAHORE UNIVERSITY OF MANAGEMENT SCIENCES**, Lahore, PK

2016-2018

**Biomedical Informatics Research Lab**

**Team Lead (2018)**

- Initiated Phase II of Theatre for in silico Oncology (TISON) – a multiscale cancer systems modeling and simulation web platform
- Worked with biologists towards incorporating more features to enhance the overall desirability of the software platform
- Developed and revamped three out of seven scales (Cell circuit design, Phenotype design, Tissue design) of the platform
- Implemented the simulation engine of the platform by integrating data from all the scales
- Trained a new team of four fresh undergraduate students and accelerated the development process by dividing scales among them

**Research Associate (2017)**

- Contributed towards the development of MATLAB-based Attractor Landscape Analysis Toolbox for Cell Fate Discovery and Reprogramming (ATLANTIS) and later enhanced the performance of the toolbox
- Developed the manuscript for the toolbox with my team and independently worked on supplementary material and user manual for the toolbox and later executed the rebuttal process with my team and dealt with a major revision within 30 days

**Research Intern (2016)**

- Gained proficiency in software development over the summer and familiarized myself with the latest trends and practices in the world of web development
- Initiated Phase I of Theatre for in silico Oncology (TISON) – a multiscale cancer systems modeling and simulation web platform and worked on cell circuit design scale to familiarize myself with the underlying software practices and later broadening my knowledge about the biology of cell cycles

**UNIVERSITY OF ENGINEERING AND TECHNOLOGY, Lahore, PK**

2015-2016

**Research Intern (2016)**

- Worked on a Linux based platform (Ubuntu) towards the development of Next Generation Sequencing (NGS) data analysis pipeline

**TEACHING ASSISTANT EXPERIENCE****DUKE UNIVERSITY, Durham, NC**

Fall 2020

- Course: Applied Biostatistical Methods I

**UNIVERSITY OF ENGINEERING AND TECHNOLOGY, Lahore, PK**

Fall 2015

- Course: Applied Biology

**PUBLICATIONS****JOURNALS**

W. McKleroy, T. Shing, W. H. Anderson, M. Arjomandi, **H. A. Awan**, et al., “Longitudinal Follow-Up of Participants With Tobacco Exposure and Preserved Spirometry.” *JAMA*. 2023; 330(5):442-453.

O. S. Shah, M. F. A. Chaudhary, **H. A. Awan**, F. Fatima, Z. Arshad, B. Amina, M. Ahmed, et al., “ATLANTIS – Attractor Landscape Analysis Toolbox for Cell Fate Discovery and Reprogramming.” *Nature Scientific Reports* 8, no. 1 (2018): 3554.

M. N. Gondal, A. Rehman, A. Arif, M. U. Sultan, **H. A. Awan**, Z. Arshad, M. F. A. Chaudhary et al., “TISON: A Next-generation Multi-scale Modeling Theatre for in silico Systems Oncology”, *Molecular Systems Biology*. (*In Preparation*)

**CONFERENCES & POSTER PRESENTATIONS**

**H. A. Awan**, M. F. A. Chaudhary, S. E. Gerard, A. P. Comellas et al., “Deep residual convolutional network predicts future severe exacerbations of COPD in SPIROMICS”, American Thoracic Society (ATS) Conference, May 2023, Washington DC, USA.

**H. A. Awan**, M. F. A. Chaudhary, W. Labaki, J. Guo, et al., “Novel QCT textures are associated with sRAGE and IL-6 in SPIROMICS”, European Respiratory Society (ERS) International Congress, September 2022, Barcelona, Spain.

**H. A. Awan**, C. J. L. Angel, G. Tomaras, “Investigating the impact of preexisting CMV immunity on the risk of HIV acquisition in HVTN505 vaccine efficacy trial for HIV”, *Duke Surgery 5th Annual Research Day*, April. 2021, Durham, NC, USA. (*abstract selected for oral presentation*)

**H. A. Awan**, C. J. L. Angel, J. McCarthy, G. Tomaras, “Preexisting CMV immunity is associated with blunted vaccine specific cellular and humoral responses and increased risk of HIV acquisition in phase IIb human vaccine efficacy trial”, *16th Annual Duke Center for AIDS Research Virtual Fall Scientific Retreat*, Sept. 2020, Durham, NC, USA. (*abstract selected for oral presentation*)

S. U. Chaudhary, M. N. Gondal, **H. A. Awan**, A. Rehman, A. Arif, R. Hussain, H. Khawar, Z. Arshad, M. F. A. Chaudhary et al., “A Next Generation Multi-scale Modeling Theatre for In Silico Oncology”, *14th International Conference on Systems Biology (ICSB 2020)*, Jan. 2020, Zurich, Switzerland.

M. F. A. Chaudhary, **H. A. Awan**, H. Hameed, Z. Arshad, M. Abubakar, K. M. Hasan, B. Wajid, et al., “TISON - Theatre for in silico Oncology: A Next-Generation Multiscale Modeling Platform for Predicting Cancer Growth, Development and Metastasis,” *DICE Mega Innovation and Entrepreneurship Summit*, Dec. 2016, National University of Science and Technology, Pakistan (*won second prize among 350 posters presented*)

M. F. A. Chaudhary, **H. A. Awan**, H. Hameed, Z. Arshad, M. Abubakar, K. M. Hasan, B. Wajid, et al., “TISON - Theatre for in silico Oncology: A Next-Generation Multiscale Modeling Platform for Predicting Cancer Growth, Development and Metastasis,” *BioSymposium 2016*, Lahore University of Management Sciences, Lahore, Pakistan (*honorable mention at the poster presentation contest*)

E. Serpedin, Z. Arshad, H. Hameed, M. F. A. Chaudhary, **H. A. Awan**, B. Wajid, “Investigating Functional Relationships of Five Critical Genes Present in the Familial Non-Synonymous Loci with Risk of having Acute Lymphoblastic Leukemia,” *Integrative Biology Conference*, July 18-20, 2016, Berlin, Germany

**HONORS & AWARDS**

- Awarded the prestigious ***Carver Fellowship*** to pursue PhD in Biomedical Engineering at The University of Iowa
- Awarded best master's project award (data science category) by the Department of Biostatistics and Bioinformatics at Duke University
- Awarded the prestigious ***Fulbright Scholarship*** to study Biostatistics at Duke University
- Granted 2nd prize for TISON - Theatre for in silico Oncology among 350 other projects at the DICE Mega Innovation Summit
- Awarded ***Student Research Grant*** for the project, "Diagnostic Tests for Cancer (now TISON)", by the National Grassroots ICT Research Initiative
- Awarded Dean's Honor Roll for Fall 2014, Spring 2015, Fall 2016, Spring 2017, Department of Electrical Engineering, University of Engineering and Technology, Lahore, Pakistan

# MARIA C ARCE

Elizabeth N.J 07202

[Carolinaarce80@hotmail.com](mailto:Carolinaarce80@hotmail.com)

(908) 397-2897

## Skills Summary

Administrative skills, communication skills, bilingual, strong customer service skill, ability to follow instructions, strong organizational skills.

## Education

***Elizabeth High School 2005***

## Certifications

NJ Dental Assistant Certificate

X-rays license

## Experience

***Teva Pharmaceutical***

Operator February 2023 to Present

disassemble and reassemble equipment, follow all standard operating procedures, perform in process checks until batch is completed, ability to demonstrate success in troubleshooting.

***Park Place Dental Group***

Dental Assistant January 2011 to Now

Preparing patients for surgery, performing administrative tasks, preparing examination areas, answering phones, verify insurance benefits.

***Amazing Smiles***

Front Desk August 2007 to December 2010

Greeting patients as they arrive, schedule and cancelling patients appointments, completing and filling insurance forms and dental billing records.

# Fengshuo Song

(530)364-7240 | songfengshuo@hotmail.com | [LinkedIn](#)

## EDUCATION

<b>University of California, Irvine</b>	Irvine, CA
<i>PhD in Computer Science</i>	Sep 2023 -
• Systems and Machine Learning, Modern Computer Systems, Embedded AI.	
<b>Columbia University</b>	New York, NY
<i>MS in Computer Science</i>	Jan 2022 - Dec 2022
• GPA: 4.0 / 4.0.	
• NLP, Database, Cloud Computing, Speech Recognition, Reinforcement Learning, ML in Genomics.	
<b>University of California, Davis</b>	Davis, CA
<i>BS Double Major in Computer Science and Statistics</i>	Sep 2018 - Dec 2021
• GPA: 3.92 / 4.0.	
• Machine Learning, Algorithm, Artificial Intelligence, Software Engineering, Design Patterns.	
• Outstanding Performance Citation in Statistics; 6 times appearance in the Dean's Honors List.	

## WORK EXPERIENCE

<b>University of California, Irvine</b>	Irvine, CA
<i>Graduate Teaching Assistant (ICS 33 Intermediate Programming)</i>	Sep 2023 – Dec 2023
• Assisted professor by administrating exams, grading assignments and exams for over 240 students.	
• Provided timely feedbacks to the questions from students.	
• Conducted lab sessions helping students with their homework assignments and projects.	
<b>Boutique Air, Inc.</b>	San Francisco, CA
<i>Full-Stack Engineer (Full Time)</i>	Apr 2023 - Jul 2023
• Migrated database from an existing system to the company's own platform, resulting in improved system integration and annual cost savings of \$250,000 for the company.	
• Improved previous system by adding new features allowing for better tracking of maintenance items of aircrafts and reducing the work for the mechanics.	
• Conducted thorough testing and debugging of programs to ensure accurate results and reliable performance.	

## PROJECTS

<b>Audio Signal Keyword Spotting System</b>	Sep 2022 - Dec 2022
• Developed using Python and Kaldi framework to detect audio commands from user.	
• Improved previous state-of-the-art results by reducing model complexity and number of operations.	
• Constructed a transformer encoder based model and trained model on MFCC features from audio samples.	
<b>Item Title Tagging System</b>	Jul 2022 – Sep 2022
• Assisted eBay manipulate user input for searching items, by assigning each query tokens a tag (i.e., brand, size, etc.), which is useful for seeking the correct items the user has requested.	
• Helped store the information for listing items by converting unstructured text data to structured data.	
• The system is built using the BERT encoder followed by a linear layer to perform tag classification.	
<b>Music Melody Synthesis Application</b>	Sep 2020 - Dec 2020
• Built a web application to generate music melodies according to user's speciation (i.e., fast, slow, strong, mild, etc.) to assisted people to compose music.	
• Constructed an Encoder-Decoder LSTM model. A kNN dimension reduction technique is used before training data.	
• The front-end uses React; The back-end uses Python Flask framework. The application is deployed on AWS.	

## LANGUAGE AND IT SKILLS

- Python, C/C++, Java, HTML/CSS/JavaScript, R, MATLAB, Go, Haskell, Lisp, Prolog.
- PyTorch, TensorFlow, Hugging Face, XGBoost, DGL, jQuery, React, Vue, Django, Flask, Docker.
- Linux/Unix, GCP, AWS, DBMS.

# CHARLES TORRES

+1 (520) 238-1686 ◊ Irvine, CA  
[torrescj92@gmail.com](mailto:torrescj92@gmail.com) ◊ [GitHub](#)

## OBJECTIVE

Improve skills in an applied context. Learn industry practices. Gain experience working in industry.

## BIO

Charles Torres is a researcher in the field of Natural Language Processing (NLP) and Machine Learning, currently pursuing his Ph.D. in Language Science from the University of California, Irvine. Charles's research focuses on language, algorithmic complexity, and learning biases. 3+ years programming experience, 2+ years Machine Learning experience.

## SKILLS

### Experience

Training DNNs • Training RNNs • Finetuning LLMs • Adapting papers • Writing findings • Presenting work

### Technical Skills

Python • PyTorch • NumPy • Pandas • C++ • R • dplyr • Julia •  $\text{\LaTeX}$  • git • PyCharm • VSCode • RStudio

### Soft Skills

Collaboration • Time management • Maintaining documentation • Detail oriented • Analogical reasoning • Problem Solving

### Other Skills

Research • Technical writing • Information theory • Statistics

## EDUCATION

### PhD Language Science

University of California, Irvine      Expected 2025

**Focus:** Algorithmic Complexity, Learning Biases

**Relevant Coursework:** Information Theory, Probabilistic Modeling

### B.Sc. Neuroscience and Cognitive Science

Minor in Mathematics      Summa Cum Laude  
University of Arizona      2017 - 2020

**Relevant Coursework:** Linear Algebra, Machine Learning, Discrete Math, Natural Language Processing, Databases

## CURRENT PROJECTS

**Neural Networks and Algorithmic Complexity.** Applied the  $L_0$ -Regularizer from Louizos et al. (2017) to PyTorch neural networks. Successfully trained LSTMs on different formal languages, measuring the accuracy and number of parameters. Presented findings at notable conferences, with plans for further extensions.

**Bio-inspired Spiking Networks.** Extending complexity measures to biologically-inspired networks. Adapted Izhikevich (2003) model into the Julia programming language.

**Linguistic NK Modeling.** Reimagined NK Models (Kauffman and Weinberger, 1989) for linguistic applications. To be used with encoder-only transformers to measure epistatic interactions in natural language.

## PAPERS

**Simpler Neural Networks Prefer Subregular Languages.** Co-authored with Richard Futrell, published in the findings of Empirical Methods in Natural Language Processing 2023. Detailed how subregular languages are learnable by simpler neural networks. Highlighted the adaptation of  $L_0$ -regularization from Louizos et al. (2017).

**Modeling Island Effects with Probabilistic Tier-based Strictly Local Grammars over Trees.** Joint work with Kenneth Hanson, Thomas Graf, and Connor Mayer. Published in the Proceedings of the Society for Computation in Linguistics 2023. Showcased that subregular formalisms combined with probabilistic projections of trees can model Wh-island constraints.

**Evaluating a Century of Progress on the Cognitive Science of Adjective Ordering.** Collaboration with William Dyer, Greg Scontras, and Richard Futrell. Published in the Transactions of the Association for Computational Linguistics 2023. Surveyed various adjective ordering theories and trained models based on measures in the cognitive literature.

## CONFERENCE PRESENTATIONS

---

**L<sub>0</sub>-Regularization Induces Subregular Biases in LSTMs.** Presented a poster at the Society for Computation in Linguistics 2023. Showcased how a simplicity preference in LSTMs using  $L_0$ -regularization enhances the acquisition of subregular patterns.

**Modeling Island Effects in Linguistics.** Delivered a talk at the Society for Computation in Linguistics 2023. Discussed the use of probabilistic tier-based strictly local grammars over trees to model Wh-island dependency judgment scores.

**Formal Languages and Computational Complexity.** Displayed a poster at Human Sentence Processing 2023. Demonstrated a human-attested language pattern learned preferentially by simpler LSTMs using  $L_0$ -regularization.

**Dependency Length Minimization.** Presented a poster at Architectures and Mechanisms for Language Processing 2020. Shared early findings suggesting dependency length minimization preferences are influenced by a speaker's native language.

**Writing Systems and Morpho-Syntactic Complexity.** Prepared a workshop presentation for the Evolution of Language Conference 2020, showcasing data on the impact of writing systems on morpho-syntactic complexity. The presentation was canceled due to Covid-19.

## OTHER PRESENTATIONS

---

**Subregularity and Computational Complexity.** Delivered a talk for the UC Irvine Language Science Community, summarizing my approach using algorithmic complexity. Emphasized that computational formalisms have innate inductive biases which can influence learning.

**Complexity Methods in Data Science.** Presented a talk for the Data Science Club. Discussed two machine learning methods for complexity measurement: one using an information bottleneck and another based on algorithmic complexity.

**Transducers and Weak Determinism.** Conducted a special lecture in Formal Language Theory. Discussed transducers and the formulation of "weak determinism" to capture certain string transformation restrictions.

## EXPERIENCE

---

### Teaching Assistant & Graduate Researcher

UC Irvine

Oct 2020 - Present  
Irvine, CA

- Taught core linguistics concepts, elevating student comprehension and performance.
- Led discussion sections, enhancing student engagement and understanding.
- Developed and adapted models from academic papers using PyTorch, Python, and Julia, driving forward department research.

### IT Help Desk

Banner Health

Mar 2016 – Feb 2020  
Phoenix, AZ

- Delivered solutions for diagnosed software problems, ensuring smooth workflow for medical professionals.
- Documented and maintained essential reference materials, aiding in faster issue resolution.

### Support Specialist

Prime Therapeutics

Dec 2014 - Aug 2015  
Tucson, AZ

- Managed call escalations, ensuring client satisfaction and trust.
- Assisted in training new employees, improving team efficiency and knowledge base.

**VAISHNAVI MEKA**  
(315) 418-3917 • vameka@syr.edu • linkedin.com/in/vaishnavimeka

---

## EDUCATION

Syracuse University, School of Information Studies, Syracuse, NY <b>M.S. Applied Data Science</b> Relevant Coursework: Database Management, Machine Learning, Quantitative Reasoning & Statistics, Introduction to Data Science	May 2025
Jawaharlal Nehru Technological University, Sri Indu Institute of Engineering and Technology, Hyderabad, India <b>B.Tech. Computer Science and Engineering</b> Relevant Coursework: DBMS, Computer Oriented Statistical methods, Design and Analysis of Algorithms, Cloud Computing, Artificial Intelligence, Data Mining, C Programming	July 2022

## WORK EXPERIENCE

<b>AI/ML Intern</b> , Mindtree, India	Mar 2022 – June 2022
• Achieved 60% process improvement through Python-based product identification filter, streamlining downstream team workflows effectively	
• Elevated user experience with insurance models by pioneering AWS-based A/B and AB tests, leveraging statistical measures for informed enhancements	
• Optimized workflows by constructing custom XGBoost images, crafting 3 pipelines, and authoring 15 Python-powered Airflow scripts to refine process flow	
• Reduced billable time by 10% and optimized costs for extensive data operations by introducing tailored Python-driven image	
• Spearheaded a 60% process enhancement by implementing a Python-based product identification filter, facilitating seamless downstream team collaboration	

## PROJECTS

<b>Crypto jacking Detection Challenge</b> - <i>Python, Stats, Machine Learning</i>	Mar 2023 – Apr 2023
• Predicted a networking activity using Machine Learning to enable better security to a user	
• Feature Engineered 18 features which helped in better identification of bad network activity	
• Improved F1 Score by 58% in identifying bad network activities by using machine learning techniques	
<b>Analyzing Motor Vehicle Accidents</b> – <i>Tableau, Dashboard, Story, Data Engineering</i>	Feb 2023 – Mar 2023
• Built a ETL pipeline for data transformations using Azure ML and Azure Databricks	
• Performed data transformations on 30GB of data using Spark and connected o/p data to Tableau using connector	
• Performed data analysis to analyze the root cause of accidents in NYC and built a story dashboard using Tableau	
<b>Facebook Friend Case Study</b> - <i>Python, Machine Learning, EDA, Visualization, Statistics, Graph</i>	Mar 2022 – Apr 2022
• Assessed potential link between two users to enable better user experience by using machine learning	
• Performed feature engineering to create 52 features for each user in a graph and modelled best fit algorithm using Python	
• Improved F1 Score by 7% in identifying link between two user nodes in a graph network using machine learning	
<b>Crime Rate Prediction</b> - <i>Python, Machine Learning</i>	May 2022 – June 2022
• Developed a web app for crime pattern prediction, employing advanced methodologies to enhance law enforcement strategies	
• Identified South Dallas as a high-crime area through rigorous hypothesis tests, informing targeted intervention efforts	
• Achieved a 78% accurate crime prediction model using KNN technique with an F1 score of 0.90	

## TECHNICAL SKILLS/CERTIFICATIONS

<b>Programming Languages:</b> Python, R, SQL	
<b>Certifications:</b> IBM Data Science Professional, Programming for everybody by Coursera	
<b>Frameworks:</b> ML Flow, Flask, DVC, GIT , TensorFlow, API, Apache Spark	
<b>Database Management/ETL:</b> SQL, Mongo DB, NoSQL, MS SQL Server, Spark	
<b>Machine Learning:</b> Regression, Forest, XGBoost, Bayesian Analysis, Decision Trees, Clustering, LightGBM, SVM, Neural Network	
<b>Statistical Knowledge:</b> T-Test, Hypothesis Testing, Confidence Interval, Correlation Analysis	
<b>Packages:</b> NumPy, Pandas, SciPy, Matplotlib, Seaborn, Scikit-Learn, PyTables, OpenCV	
<b>Software Tools:</b> Pycharm, VSCode, MSEExcel, MS Powerpoint, MS Word, Tableau, Gitlab, Jupyter Notebook, R Studio	

# Owen Grimm

owen.z.grimmm@vanderbilt.edu ♦ (314)-813-3103 ♦ Nashville, TN ♦ [LinkedIn](#)

**Skills:** Python, C++, Java, JavaScript, HTML, CSS, SQL, & R; OpenAI API; TensorFlow; Pytorch; AI Training; Microsoft 365; Adobe; Data Wrangling; Website Designing; Marketing; Revenue Modeling; Sales; Photography

## EDUCATION

### Vanderbilt University, Nashville, TN

Expected May 2025

Bachelor's Degree, Majors: Computer Science & Neuroscience, Minor: Mathematics

- **Cumulative GPA:** 3.7/4.0
- **Awards & Honors:** Cornelius Vanderbilt Scholar, Dean's List
- **Relevant Coursework:** Software Design, Data Structures, Programming & Problem Solving, Computational Expression, Algorithms, Discrete Structures, Machine Learning, Data Science, Probability & Statistics, Calculus
- **Activities:** Google Developer Student Club, Data Science Club, Code Ignite, Innovation & Entrepreneurship Society, Men's Club Soccer, Vanderbilt Pre-Medical Society, Project C.U.R.E, Sigma Nu Fraternity (Risk Chair)

### Allegheny College, Meadville, PA

Aug. 2021 - May 2022

Bachelor's Degree, Majors: Computer Science & Biology

- **GPA:** 4.0/4.0
- **Awards & Honors:** Distinguished Alden Scholar (GPA > 3.8), Trustee Scholarship Recipient, Lambda Sigma Honors Society, 2021 Men's Soccer Most Improved Award, Global Citizen Scholarship
- **Activities:** Men's Varsity Soccer, Pre-Health Club, Health Coaches Program, Global Citizen Scholars

## PROFESSIONAL EXPERIENCE

### Presskit AI

August 2023 – Present

Software Engineer

New York, NY

- Built and maintained software applications for media contact database and AI-driven writing tool
- Leveraged Python, JavaScript, React, and modern web technologies to enhance overall functionality of software
- Worked directly with Chief Technology Officer to implement new design and testing features

### Martek Global Services, Inc.

June 2023 – August 2023

AI Software Engineer

Bethesda, MD

- Integrated AI-driven projects into operations, yielding workflow optimization and tenfold increase in efficiency
- Consulted with a multidisciplinary team to drive operational enhancements and spearhead innovative solutions
- Led the enterprise-wide integration of Microsoft's Azure Cloud Computing Platform and Monday.com

### Johns Hopkins University Hospital

June 2022 - August 2022

Cardiac Physician Apprentice

Baltimore, MD

- Shadowed 50+ hours of cardiac surgery including coronary artery bypass grafting and aortic valve replacement
- Developed an understanding of the critical role of effective communication among diverse specialists in an intense, operating room environment

## PROJECTS

### PDF Interaction AI Bot – Python, Streamlit, & OpenAI API

- Built full-stack web app using Streamlit UI to allow users to upload and query pdf documents
- Utilized OpenAI API in Python to access LLM for generative outputs in response to user queries
- Managed integration across all operational facets at Martek, yielding a 95% improvement in project efficiency

### Email Genius – VS Code & OpenAI API [[GitHub](#)]

- Developed an auto-GPT utilizing VS Code and OpenAI's API to create an application that provides tailored company recommendations and talent acquisition email content

### Fine-Tuned Fantasy Football LLM – GPT-3, Kaggle, & Excel

- Generated fine-tuned GPT-3 model leveraging Excel and Kaggle dataset of NFL player performance statistics

# Tianmin Zhao,

Phone: (864)986-7552 | Email: [tianmiz821@gmail.com](mailto:tianmiz821@gmail.com) | LinkedIn: [/in/tianmin-zhao-25b972244](https://in/tianmin-zhao-25b972244)

## SUMMARY

PhD candidate with 3+ years experience in marketing industry, using linear Lasso and logistic estimating advertising Demand Boost, extracting insights and providing suggestions for companies with asymmetric advertising information.

## SKILLS

Python, Scikit-Learn, Seaborn, Numpy, Scipy, Pandas, Matplotlib, Spark, Tensorflow, Keras, Flask, Git, Github, GCP, SQL, Linux/Unix, selenium, Requests, R, Causal Inference Analysis, latex

## EDUCATION

<b>Clemson University</b>	Clemson, SC
<i>Ph.D in Economics</i>	08/2016 - 12/2023
<i>Master in Mathematics</i>	08/2021 - 05/2023

## EXPERIENCE

<b>Clemson University</b>	Clemson, SC
<b>Data Scientist/ Research Assistant</b>	08/2018 - 12/2022

### ***Twitter Politician argue Insight***

- To identify the causal relation between Russian twitter post to result of election, collected and investigate twitter accounts data who reported by US government.
- To gain and clear our dataset, our group of research assistant meet and cooperate with each other weekly after downloads 100 million twitter account from API.
- Find and report some statistical difference before and after the Election as evidence that Russian troll twitter account behave differently.
- To identify the political standpoint, build NLP Sentiment Analysis model. Label part of twitter by hand and train a NLP model.

### ***Tiktok Topics Popularity Estimator***

- To help new Influencers understand social media Topics' popularity changes, lowering their entry cost of this new industry. Built an Hawkes Process Kernel Estimator based on Hawkes process model.
- I scraped Tiktok top 100,000 influencer's video from 2020-2021, clean and train a model with different linear combination of distribution family.

## PROJECTS

<b>Advertising Influencer Detector</b>	06/2023 - 08/2023
----------------------------------------	-------------------

- To help Marketing representatives effectively identify the influencers who will boosting their sales and reducing advertising budget, built an Influencer Detector based on binary classification models.
- To gain training data, built web scraping of TikTok for influencer information and Taobao website for sales history; trained and optimized multiple models (Logistic Regression, Random Forest, and Xgboost).
- Found Xgboost is the best performing model with a Precision score of 30%, which is 50% better than the traditional trial-and-error method. Reduces 22.5% advertising budget wasting.

# Meet Panjwani (MPH, BDS)

(857)-308-6058 | meet1995@bu.edu | <https://www.linkedin.com/in/meetpanjwani/> | Boston, MA 02134.

## **ABOUT ME**

Determined Statistician inclined towards Statistical Programming, Machine Learning and Artificial Intelligence with enormous enthusiasm for solving problem which are geared for Drug Discovery.

## **EDUCATION TRAJECTORY**

### **Master's in Public Health (Epidemiology and Biostatistics)**

Boston University School of Public Health, Boston, MA. [GPA 3.68]

### **Bachelor of Dental Surgery**

Gujarat University, Gujarat, India. [GPA 3.5]

## **TECHNICAL ACUMEN**

- R, SAS (base and macro), Python, Linux
- SQL server, MySQL, Redcap
- Docker, Git
- Microsoft Office (Excel, Word, PowerPoint, Teams)

- Array Studio, Ingenuity Pathway Analysis, Bioturing
- Python (Matplotlib, numpy, pandas, scipy, scikitlearn, nltk)
- R (tidyverse, Bioconductor, Shiny, Markdown, caret, tensorflow)
- SAS (Base, Macro, graph, ODS)

## **TECHNICAL SKILLS**

- Regression (Linear, logistic, cox, nominal, polynomial, Non-Parametric)
- Mixed Models, Survival analyses, GLM,
- ANOVA, ANCOVA, Mixed models
- Database Management, Data modelling and Data Harmonization
- Sample Size and Power calculations
- Statistical Analyses Plans,
- Interim Analyses and DSMB
- Tables, Listings and Figures

- Quality Control, Quality Assurance, SOPs, DMP, SDLC
- Dimension Reduction (PCA and EFA)
- Descriptive and inferential analyses
- Arrays, loops, Macros, IML, SQL, format, Retain
- Case Report forms
- Simulations, Adverse events, shift tables
- Real world Data (EHR, Clinical trials, Survey, Multi-omics)
- Imputations and propensity score adjustment
- CDISC (SDTM and Adam) & i2b2
- NGS Pipelines (RNA Seq analysis, Single Cell Analysis)

## **PROJECTS AND TASKS**

- International Stroke Trial (19500 subjects): Examined effects of Aspirin on mortality for stroke patients using survival analysis (**SAS, R and Excel**)
- BMS stent vs DES stent (750 subjects): Superiority clinical trial analyses. Performed multiple imputation and propensity score adjustment and analysed multiple end points (**SAS**)
- Statistical modelling (700 subjects): Selected best model using backward selection based on adjusted Rsq to predict Z scores of birth weight. (**Excel, SAS**)
- ENACT study (3.5 million patients): Extracted cases and controls and their clinical data from SQL server database (i2b2 model) in Rstudio using RJDBC connection, mapped the diagnosis codes to CCSR categories to assess effects of long covid. (**SQL, R**)
- Covid-19 forecasting (150k patients): Predicted mortality and hospitalization rates at MGB network using 8 forecasting models and made a single function that selected best model based on RMSE value. (**R,SQL**)
- CD28 OX 40 MLR Transcriptomic study: - Compared pathways among various samples by creating NGS data pipelines, followed by data visualization and reporting results (**Array Studio, R Studio (Bioconductor Packages), Ingenuity Pathway analysis**)

## **CAREER NARRATIVE AND MILESTONES**

- **Research Assistant, BOSTON UNIVERSITY SCHOOL OF PUBLIC HEALTH [PRESENT]**
  - Analysing prompts from various AI chatbots and comparing them with human judgement.
  - Utilizing NLP libraries across python and R studio for sentient analysis and text mining.
  - EHR based Phenotype code mapping to diseases in ACT database (i2b2 model) using medical terminologies like ICD, CPT, LOINC, NDC etc.
  - Performing predictive and Descriptive analysis.
  - Interpreting the results and presenting them using interactive visualizations from r packages like Plotly and dygraph.
  - Database management, Data cleaning, Data harmonization.
- **GENOMICS DATA SCIENCE CO-OP, SANOFI, BOSTON, MA.** Aug'22-Dec'22
  - Applying statistical models on multi-omics data for biomarker identification and patient stratification.
  - Using computational tools to perform analysis on Multi omics data.
  - Building NGS data pipelines.
  - Utilizing Machine Learning Approaches for Multi omics data analysis (GLE, Dimension Reduction, Exploratory Data Analysis, Supervised Learning, Unsupervised Learning, Statistical modelling and cluster analysis)
- **RESEARCH ASSISTANT, BOSTON UNIVERSITY SCHOOL OF PUBLIC HEALTH, BOSTON, MA.** May'21-May'22
  - Data extraction from MICS (survey data) database, data cleaning and data harmonization in R.
  - Performing statistical modelling, statistical analyses, visualization, and report writing for both technical and non-technical audiences.
  - Conducted a comprehensive literature review on the risk factors, prevention, and treatment of diarrheal diseases in underdeveloped countries.
- **GRADUATE TEACHING ASSISTANT, BOSTON UNIVERSITY SCHOOL OF PUBLIC HEALTH BOSTON, MA.** Jan'22-May'22
  - Hosted one on one TA session every week to review lectures and reviewed lectures with class students for Introduction to Biostatistics class (SAS).
  - Assisted the professor in grading of assignments and exams.
- **RESEARCH ASSOCIATE, NIRMIT CLINICAL RESEARCH, AHMEDABAD, INDIA** Sep'17-Dec'20
  - Performed general dental procedures and Qualitative analysis for selection of patients in experimental study.
  - Assisting in designing, conducting, and collecting clinical trial data and their analyses in SAS.
  - Responsible for Data Entry via the clinic's online tool and assisted the Client Services team on support issues and other issues as needed.
  - Patient Recruitment, Clinical trial site procurement, Coordination with PIs and CROs.

# Lynne Auger

(484) 634-1347 | auger.lynnne@gmail.com | Danielsville, PA

Highly motivated data analyst-student seeking a challenging position to implement and expand current analytical skills. Experienced working with Python, and R to investigate for patterns revealing insights through storytelling and visualizations. Ability to work independently as well as in a team environment.

## EDUCATION:

**Western Governors University** | Salt Lake City, UT

7/2024

BS Data Management and Data Analytics

Courses:

- Data Analytics
- R Analysis
- Machine Learning
- Data Wrangling
- Data Visualization

## CERTIFICATIONS:

CompTIA A+

6/2022

Network+

8/2022

Project +

4/2023

Udacity nanodegree Data Science (in progress)

## INFORMATION TECHNOLOGY COMPETENCIES:

Systems: Windows PC

Languages: Python, MySQL, Java, R

Software: MS Word/Excel/PowerPoint

## WORK ETHIC PROFILE:

Team Player, Communication, Reliable, Consistent, Diligent

## MEMBERSHIP:

**The National Society of Leadership and Success (NSLS)**- Member: Nov 2022 – present

Selected by campus administration to participate among the top students in a leadership program.

## PROJECTS:

**Title: Data Cleaning**

**Tools used: Python, R**

Collected semi-structured data, organized and checked it for completeness, validity, and uniqueness. Cleaned the data to be consistent throughout the file to minimize errors during the analysis used to uncover new and pertinent insights.

**Title: Bikeshare Analysis**

**Tools used: R**

Performed an analysis of Bikeshare services to determine the trends in usage across three major cities using rental times, durations and days of the week. This analysis uncovered trends that suggest the service is widely used for short durations such as transportation to and from work and school.

**Title: Data Cleaning**

**Tools used: Python, Pandas, Numpy**

Analyzed the impact of Genre on the financial success of movies. Compared the popularity ratings to the revenue uncovering no correlation. Prepared box plots to view the revenue trend finding that Action movies have the largest variation.

# Jiayu Li

[linkedin.com/in/jiayu-danielle-li-158102188/](https://linkedin.com/in/jiayu-danielle-li-158102188/) • jiayu7@uw.edu • (206) 965-0677

## SUMMARY

Passionate computer science Ph.D. student with practical experience in implementing machine learning and Computer Vision solutions. Seeking an internship opportunity to leverage technical skills and creative problem-solving ability to contribute effectively, acquire valuable industrial experience, and further enrich my understanding of the research.

## EDUCATION

University of Washington			US
Computer Science and Systems	Ph.D.	12/2020 - present	
University of Washington			US
Computer Science and Systems	GPA:3.99	Master	09/2019 - 12/2020
Beijing University of Posts and Telecommunications			Beijing, China
Computer Engineering	Bachelor	09/2015- 07/2019	

## PROFESSIONAL EXPERIENCE

Roadside Parked Vehicle Classification	Python	03/2023 - present
• Investigated and applied map matching algorithm to cut videos aligned with road segments using OpenStreetMap data.		
• Implemented cutting-edge computer vision techniques like Grounding DINO for zero-shot vehicle detection and road surface detection, and ByteTrack for efficient vehicle ID assignment.		
• Researched AutoML, and built a baseline classifier to categorize parked or non-parked vehicles, providing essential statistics of roadside occupation for further analysis. Achieved overall 91% accuracy on 4-hour videos.		
Smart Street Parking Sign Understanding System	Python	08/2020 – 12/2021
• Researched 20+ state-of-the-art object detectors, including YOLO and Swin Transformer. Fine-tuned models by annotated parking sign images, and obtained 97% AP@.5 on sign detection, 99.% AP@.5 on symbol detection.		
• Optimized CDistNet by integrating customized sign text dictionary into training loss calculation and post-processing procedures. Achieved 91% accuracy of text recognition on the large dataset, outperforming Google OCR tools.		
• Conceived and created a powerful and easily extensible rule generator algorithm to extract semantic data from unstructured parking sign text. Successfully interpreted 40+ diverse parking sign text structures with 100% accuracy.		
• Implemented Django application served with Nginx and Certbot, publishing 50+ APIs for frontend collaboration.		

## WORK HISTORY

Trio.AI	Beijing, China
Machine Learning Intern	01/2019– 07/2019
• Analyzed 50k users' behavior by MySQL, filtered the accidental clicking, and formed DAU analysis reports.	
• Coded web crawlers by Selenium to gather the latest news titles and built a 10 GB corpus in ElasticSearch server.	
• Crawled 70k+ users' clicking history in baike.com search engine as training data, and fine-tuned BiLSTM model for generating sentence embeddings to predict the search intent by ElasticSearch and semantic similarity score.	
• Developed and maintained the high-performance backend system by Djano to predict the search intent, and boosted user engagement by 11%	

## PROJECT EXPERIENCE

Ariel ML Data Challenge (Regression)	Python	06/2022 - 10/2022
• Conducted the feature engineering, and applied state-of-the-art deep learning architectures, such as VGG16, ResNet, DenseNet, EfficientNet, to extract additional vector features from the spectra images for pattern recognition.		
• Performed transfer learning by pre-trained backbone with 50k planets data. Applied the multi-target regression model with learning rate decay and early stopping, and achieved the score of 94.4% on the test dataset in PyTorch		
• Optimized neural network architecture, improving the score by 2.2%. Ensembled 15 models to get the top 10.		

Big Data Hotel Reviews Truthful or Deceptive Prediction	Tensorflow	12/2020 - 03/2021
• Performed effective data cleaning, preprocessing, and feature enhancement by PCA, word2vec, and TF-IDF analysis.		

• Trained 10+ machine learning models including Random Forest, XBG, SVM, and LSTM, and attained 93% accuracy.

## SKILLS & OTHERS

- Programming Language: PyTorch, Python, Java, SQL, JavaScript, PySpark, C++, Pandas, Tensorflow
- Others: Linux, BASH, AWS, Docker, Django, Nginx, Github, Selenium, Computer Vision, NLP

## PUBLICATIONS

- [1] Li, Jiayu, et al. "Automatic Street Parking Sign Reading." *Internet of Things, Ad Hoc and Sensor Networks Technical Committee Newsletter (IoT-AHSN TCN)* 1.14 (2021).
- [2] Chau, H., Jin, Y., Li, J., Hu, J., & Cheng, W. Real-Time Street Parking Sign Detection and Recognition. *IJCAI'22 AI4AD (Artificial Intelligence for Autonomous Driving) Workshop*, M.Sc. 2021

# MORIYAH SCHICK

516-301-7633 | mschick@gradcenter.cuny.edu | Moriyah Schick | moriyahschick

## Education

**The Graduate Center, City University of New York, NY, NY** | Expected May 2025

*PhD in Computer Science*

GPA: 3.87/4.00

**Coursework:** Advanced NLP, Machine Learning, Artificial Intelligence, Big Data Analytics

**Languages:** Python (Advanced), Java, Bash, SQL

**Yeshiva University, NY, NY** | May 2020

*Bachelor of Arts in Computer Science; Minor: Mathematics*

GPA: 3.90/4.00

**Awards:** Summa Cum Laude, Computer Science Department Award for Excellence in Computer Science 2020

## Research Experience

**CUNY Graduate Center**

Sept 2020 – Present

*Graduate Research Assistant*

NY, NY

- Researching grammatical error correction for low resource languages (Arabic, Ukrainian, Russian, Spanish)
- Testing different approaches including using transformer models and MT5
- Creating synthetic data for pretraining neural models in addition to use of large language models

**Yeshiva University**

Sept 2019 – May 2020

*Research Assistant*

NY, NY

- Researched leading words in a project to detect leitwort (“leading words”) in the Bible in a computational manner according to Umberto Cassuto’s approach
- Worked on senior thesis, “Genre Analysis via Constituent Tree Structure.” Used Python, NLTK and Scikit-Learn.

## Work Experience

**Deloitte**

June 2022 – Aug 2022

*Machine Learning & Natural Language Processing Intern*

NY, NY

- Built a program that automatically tagged financial statements with appropriate XBRL tags
- Created a custom scraper to get data, preprocessed appropriately, and trained models using FastText
- Final accuracy rates achieved were above 80%

**Schonfeld**

June 2021 – Aug 2021

*Quantitative Research Intern*

NY, NY

- Researched applying NLP techniques to identify alphas for implementation in a black box model
- Preprocessed datasets, finetuned machine learning models to test my hypotheses

**Citigroup**

June 2019 – Aug 2019

*Summer Technology Analyst*

NY, NY

- Implemented a program that automated workflow; utilized Apache Airflow, Python, Linux, Bash, Flask, and SQL

## Teaching Experience

**CUNY Graduate Center**

Fall 2021 – Present

*Teacher's Assistant*

NY, NY

- Advanced NLP – Spring 2023
- Machine Learning – Spring 2022, Spring 2023, Fall 2023
- Intro to NLP – Fall 2021

**Yeshiva University**

May 2023 – Present

*Thesis Mentor*

NY, NY

## Projects

**NYC Crime Prediction Project** | *Pandas, scikit-learn, NumPy, Matplotlib, Git*

Feb – May 2021

- Used historical crime complaint data for New York City to classify 3 crime types and predict the weekly and daily average count of 4 crime types
- Steps taken include accounting for bias, feature selection and engineering, dimensionality reduction, missing data clean up, and considering the explainability of our models and results

**Spam Detection Class Project** | *Pandas, scikit-learn, NumPy, Matplotlib*

Month – Month Year

- Implemented supervised and unsupervised ML algorithms to detect spam text messages (naïve bayes, k-means clustering, logistic regression, SVM, multi-layer perceptron), used bag of words and TF-IDF representations
- Accuracy rates ranged between 88-99.3%

## Publication

**Leitwort Detection, Quantification and Discernment** | *Publication link*

Jan 2020

# XINWEI HUANG

xinweihuang2024@gmail.com

716 650 8482

## OBJECTIVE

---

Seeking an intern Statistician/Data scientist position in biotechnology/pharmaceutical industries where I can leverage my expertise to conduct in-depth statistical analysis using real-world data and develop innovative methodologies to address emerging challenges.

**Skills:** Statistical analysis, Statistical inference, R programming, Statistical software development, Statistical simulation, Computational statistics, Bayesian analysis, Signal detection, Machine learning, CART, Random forest, Regression, Copula, Survival analysis, Time series, Bootstrap

## EDUCATION

---

<b>Ph.D. candidate in Biostatistics</b> (State University of New York at Buffalo)	<b>Aug 2021 - present</b>
<b>Ph.D. student in Statistics</b> (National Chiao Tung University)	<b>Sep 2019 - Jul 2021</b>
<b>M.S in Statistics</b> (National Central University)	<b>Sep 2017 - Jun 2019</b>
<b>B.B.A in Statistics and Information Science</b> (Fu Jen Catholic University)	<b>Sep 2013 - Jun 2017</b>

## WORK EXPERIENCES

---

<b>Roswell Park Comprehensive Cancer Center - Research Assistant</b>	<b>Aug 2023 - present</b>
<i>Serving as research assistant and statistician in Department of Cancer Prevention and Control. Handling data cleaning and statistical analysis in cancer research for Dr. Yao Song's team.</i>	

- **Association analysis for proinflammatory cytokines and biomarkers:** Updated cytokine data from database, ensured data accuracy and integrity. Conducted statistical analysis to investigate associations between proinflammatory cytokines and critical biomarkers, including bone turnover markers, bone regulatory markers, and bone mineral density. Utilized ANOVA tests to identify significant correlations and provide valuable insights into cancer research.
- **Meta-analysis of clinical trials involving immune checkpoint inhibitors (ICIs):** Spearheaded data cleaning and quality assurance, ensuring reliable and accurate results. Calculated Enrollment-incidence ratio (EIR) and conducted statistical estimation for minoritized ethnic groups. Applied bootstrap method to obtain confidence intervals.

<b>Xiamen Chazen Biomedical Technology Company - Intern Statistician</b>	<b>Jul 2020 - Sep 2020</b>
<i>Provided statistical support to the team as a statistician, primarily on projects related to oncology research, real world study, and phase III clinical trials of vaccines.</i>	

- **Lung cancer study:** Conducted comprehensive statistical analysis, including descriptive statistics, Kaplan-Meier, and Cox regression, to provide valuable insights into the research. Employed the Random Forest algorithm to develop a Drug Recommendation System, enhancing treatment strategies. Prepared detailed technical reports and summarized findings in conclusive reports.
- **COVID-19 research:** Collaborated with the medical research team to build a statistical model for the COVID-19 research project. Utilized the CART (Classification and Regression Trees) method to identify crucial interaction terms, contributing to a deeper understanding of the data.
- **Phase III clinical trial of HPV vaccine:** Assisted in reviewing the study protocol, ensuring compliance with best practices and industry standards.

# **RESEARCH PROJECTS**

---

## **Bayesian framework for medical product safety**

- Conducted research to determine adverse events (AEs) of concern of drugs or other medical products from spontaneous reporting system databases in pharmacovigilance.
- Proposed a formal multilevel Bayesian framework that can acknowledge the similarity and interaction between drugs and AEs more faithfully.
- An R package implementing the methods is developed based on Stan.

## **Copula-based Markov chain and its applications**

- Researched on mathematical properties of the copula-based Markov chain model. Conducted simulation studies to compare the proposed method with traditional time series model. Proposed hypotheses testing method for goodness-of-fit test.
- Applied the model to diverse datasets, including stock market data for risk management, chemical data for quality control, and sports data for anomaly detection, demonstrating its adaptability and effectiveness.
- Extended the copula-based Markov chain model to survival analysis, introducing a novel approach for handling semi-competing risks and recurrent events, with a focus on cancer data containing multiple event times and a single primary endpoint.
- Released two dedicated R packages to facilitate the utilization of these models across different applications.

# **STATISTICAL SOFTWARE DEVELOPMENT**

---

**pvBayes**: An R package implementing signal detection method for pharmacovigilance using Bayesian sampling, estimation, and inference.

**Copula.Markov.survival**: An R package for likelihood estimation and corresponding analysis under the copula-based Markov chain model for serially dependent event times with a dependent terminal event

**Copula.Markov**: An R package for performing estimation and statistical process control under copula-based time-series models.

# **PUBLICATIONS**

---

## **Articles**

### **- Medical Research**

- [2] Beshara, M., Sheng, H., **Huang, X.-W.**, Roh, J. M., Laurent, C. A., Lee, C., ... Kwan, M. L. (2023+). A Prospective Study of Vitamin D, Proinflammatory Cytokines and Risk of Fragility Fractures in Women on Aromatase Inhibitors for Breast Cancer . *Under review*.
- [1] Chua, A., Delmerico, J., Sheng, H., **Huang, X.-W.**, Liang, E., Gandhi, S., Puzanov, I., Sakoda, L. C., Morrow, G. R., Ambrosone, C. B., Kamen, C., & Yao, S. (2023+). Underrepresentation and Underreporting of Minoritized Racial and Ethnic Groups in Clinical Trials on Immune Checkpoint Inhibitors. *Under review*.

### **- Statistical Method**

- [3] **Huang, X.-W.**, & Emura, T. (2022). Computational methods for a copula-based Markov chain model with a binomial time series. *Communications in Statistics - Simulation and Computation*, 1–18. doi:10.1080/03610918.2022.2061514

- [2] **Huang, X.-W.**, Wang, W., & Emura, T. (2021). A copula-based Markov chain model for serially dependent event times with a dependent terminal event. *Japanese Journal of Statistics and Data Science*, 4(2), 917–951. doi:10.1007/s42081-020-00087-8
- [1] **Huang, X.-W.**, & Emura, T. (2021). Model Diagnostic Procedures for copula-based Markov chain models for Statistical Process Control. *Communications in Statistics - Simulation and Computation*, 50(8), 2345–2367. doi:10.1080/03610918.2019.1602647

## Books

- [1] Sun, L.-H., **Huang, X.-W.**, Alqawba, M. S., Kim, J.-M., & Emura, T. (2020). *Copula-based Markov models for time series: Parametric inference and Process Control*. Singapore: Springer.

## AWARDS AND HONORS

---

Dr. Richard N. Schmidt Scholarship	2021, 2022
The Best Thesis of the Chinese Statistical Association (Taiwan)	2019
University Scholarship for Excellent Mainland Chinese Students (NCU)	2017, 2018
The First Prize of Research Project Contest (FJU)	2016
University Scholarship for Excellent Mainland Chinese Students (FJU)	2016

## PROFESSIONAL MEMBERSHIP

---

American Statistical Association (ASA)	2023 - Present
IBS Eastern North American Region (ENAR)	2023 - Present

## SERVICES

---

Member of UB SPHHP Graduate Academic Action Committee	2023 - present
President of UB Buffalo ASA Student Chapter	2023 - present
Founder of UB Buffalo ASA Student Chapter	2023

# Malini Shivaram

(408) 896-8452 • malinish@andrew.cmu.edu • linkedin.com/in/malinishivaram

## EDUCATION

<b>Carnegie Mellon University</b> , Pittsburgh, PA <i>Master of Science in Artificial Intelligence and Innovation</i>	August 2023 – Present
<b>Carnegie Mellon University</b> , Pittsburgh, PA <i>Bachelor of Science in Artificial Intelligence</i> – University Honors, Senior Leadership Recognition Award, SCS Dean's List	August 2019 – December 2022

## SKILLS & RELEVANT COURSEWORK

**Programming Languages:** Python, C, Java, MATLAB, R, SML

**Frameworks, Libraries, & Systems:** Azure ML, AWS, Apache Spark, Pytorch, Tensorflow, OpenCV, numpy, scikit-learn, Git, Linux

**Coursework:** Machine Learning With Large Datasets, Foundations of Computational Data Science, Intro to Machine Learning, AI & Future Markets, Intro to Deep Learning, Intro to Computer Vision, Intro to Robotics, Human-AI Interaction, AI Methods for Social Good, Parallel and Sequential Data Structures and Algorithms

## EXPERIENCE

### AI Developer

January 2023 – August 2023

Stottler Henke Associates, Inc., San Mateo, CA (Remote)

- Led the development of a prototype from start to finish that automatically generates behavior models from over 10,000 textual news documents using NLP tools. Presented a company tech talk about the project.
- For another project, performed data analysis, data exploration, and machine learning experiments in Python on satellite control network data for fault detection. Employed methods such as clustering, topic modeling, and text classification using Scikit-learn, Tensorflow, and HuggingFace models.
- Wrote proposals for SLAM-based computer vision solutions that are each worth over \$250,000 of funding.

### AI Developer Intern

May 2022 – August 2022

Stottler Henke Associates, Inc., San Mateo, CA

- Coded and trained complex machine learning models using Tensorflow in Python to visually classify species of pathogenic bacteria. Systematically tuned my models to result in my best model reaching an accuracy of 93%.
- For another project, designed and implemented a detailed scheduler from scratch in Java that determines when and where bundles of data are sent through a lunar network of satellites.

### Research Assistant

June 2021 – December 2022

NORILLA, Pittsburgh, PA

NoRILLA is a patented mixed-reality educational system to improve STEM learning, co-founded by Dr. Nesra Yannier at CMU.

- Designed and implemented an application from scratch in Java using computer vision concepts such as blob detection that efficiently captures image data to be used for training feature matching computer vision systems.
- Led the implementation of new features, such as corner detection and solidity measures, to improve the system.

### AI Intern

June 2020 – January 2023

Kashmir World Foundation, Great Falls, VA (Remote)

KwF is a non-profit organization that uses AI and drones to help save endangered species.

- Trained and tuned a convolutional neural network to detect and classify types of sea turtle tracks for the MiSHELL Sea Turtle Project, resulting in a model with a 92% mean Average Precision. Also made similar contributions for KwF's Snow Leopard project with the intent of classifying different Himalayan species.
- Presented my work in webinars and panels, and started co-authoring a paper that showcases the developments and results of the sea turtle project.

## PROJECTS

### Free Body Diagram Scorer

August 2021 – December 2021

- Built a complete application in Python that scores handwritten physics free-body diagrams, using computer vision tools such as Hough transforms and a convolutional neural network.
- Incorporated human-AI interaction concepts to improve the interpretability of application outputs.

# Shuliang Yu

227e, 6 Affinity Ln,  
Buffalo, NY. 14215

716-730-1468

[mogydaomin@outlook.com](mailto:mogydaomin@outlook.com)

## Research Interests

Longitudinal data , Survival Analysis, Clinical Trials

## Education

University at Buffalo	Buffalo, NY.	Aug. 2021-May 2022 (Expected)
● MA in Biostatistics & PhD course work	Relevant course work: Theory of Statistical Inference, Theory of Linear Models, Intro to Theoretical Statistics I & II, Applied Statistics I & II, Intro Stat Inference, Advanced Statistical Computing, Categorical Data Analysis, Longitudinal Data Analysis, Survival Analysis, Calculus III (Multivariate), Linear Algebra, Study of Health Behavior, Public Health Population Well-being	

Rhodes College	Memphis, TN.	Aug. 2016-Aug. 2020
● BS in Neuroscience	● Rhodes Award for four years	● Relevant course work: Python, Java, C++, Discrete Structures, Biology I & II, Neuroscience, Molecular Neuroscience, Mechanisms of Development, Cell Biology, Neuroendocrinology, Movement Neuroscience, General Chemistry, Organic Chemistry I & II, Analytical Chemistry, Physiology & Anatomy I & II, Biochemistry, Calculus II (Integral), Physics I & II

## Experience

Project: A shrinkage method for two sample comparison on the best linear combinations of multiple biomarkers	Fall 2022
● Instructor: Dr. Albert Vexler, Department of Biostatistics, UB 1. Proposed a Stein's type estimate of the best linear combinations of multiple markers 2. Derived the large sample properties of the proposed estimate 3. Designed a bootstrap algorithm for computing p-value and confidence interval 4. Conducted Monte Carlo simulations under different settings	

## Research on MiRNAs that Regulate Ferroptosis-Related Genes

May 2021-Aug 2021
● Instructor: Hanwen Lu, PhD Candidate at Xiamen University, Center of Neurology & Neurosurgery, The First Affiliated Hospital of Xiamen University ● Cultured, froze and cryopreserved glioblastoma cell cultures for experiment ● Conducted Co-Immunoprecipitation, Western Blot, Gel Electrophoresis, and qPCR to compare between selected miRNA-regulated cells and control cells

## Research on Epigenetic Markers of Fetal Alcohol Spectrum Disorder

Fall 2019-Summer 2020
● PI: Dr. Kristin Hamre, Department of Anatomy and Neurobiology, University of Tennessee Health and Science Center (UTHSC) ● Dissected mouse embryos and performed Polymerase Chain Reaction and Liquid Chromatography to determine the sex of mouse embryos ● Assisted in drafting for the textbook <i>Springer Clinical Medicine - Neurodevelopmental Pediatrics: Genetic and Environmental Influences</i>

## Thesis: The Role of Direct Pathway of Movement in Functional Hand Paresis

Jan. 2020-May 2020

● Instructor: Dr. Tanushree Pandit, Department of Biology, Rhodes College ● Explored whether traumatic emotional experience will impact the nerve signals towards striatum and therefore hinder smooth movements
---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

## **Project: Corticosterone Did Not Affect Oxytocin Expression in the Hypothalamic Paraventricular Nucleus of a Chronic Stress Model**

Nov. 2019

- Instructor: Dr. David Kabelik, Chair of Neuroscience, Rhodes College
- Designed and applied experimental methods including Behavioral Interaction, Elevated Zero Maze, Open Field Chamber, Light-Dark Apparatus, Novel Object Test, Hyponeophagia test (HNP), Enzyme-Linked Immunoassay (ELISA)
- Conducted data analysis using Mann-Whitney U test, Principal Component Analysis and Spearman's rho, etc.

## **Research on Biophysical Underpinnings of Epilepsy**

Fall 2017-Spring 2018

- Instructor: Dr. Kelly Dougherty, Department of Neuroscience, Rhodes College
- Analyzed mice neural activities under the influence of antiepileptic drugs for subsequent modeling
- Obtained data from mice brain slices for subsequent modeling and analysis using ImageJ
- Simulated neural activities using NEURON

## **Extracurricular**

---

- Teaching Assistant for STA119: Statistical Methods, UB 2022
- Secretary of the Graduate Student Association, UB 2022
- Resident Assistant at Rhodes College Residence Life 2017-2019
- Departmental Assistant at the Math, Chemistry and Political Science Departments 2016-2020
- Piano Accompaniment for Chamber Music, Rhodes Symphony, Harp Ensemble 2018-2020
- Volunteer for the Rhodes' Food Recovery Network 2019
- Member of Rhodes Crew Team 2019

## **Technical Skills**

---

- Proficient in R, SAS
- Familiar with Python, Java

## **Standardized Test**

---

- GRE 08/31/2019 Total 329 Verbal 162 (90%) Quantitative 167(89%) AW 4.0(55%)

# THARUN PASHAM

[pashamtharun@gmail.com](mailto:pashamtharun@gmail.com) | Mobile:3462815014 | TX, 77054 | [LinkedIn](#)

## PROFESSIONAL SUMMARY

Results-driven data scientist with a proven track record of extracting meaningful insights from complex datasets and developing machine learning models. Skilled in data cleaning, statistical analysis, and data visualization, with a strong ability to communicate findings to both technical and non-technical stakeholders.

## EDUCATION

- **Master of Science in Data Science** Jan 2023-- May 2024  
**University of Houston** – Cullen College of Engineering, Houston, TX. **GPA - 3.45/4.0**  
**Coursework:** Data analysis, Advanced Machine Learning, Data Mining, Advanced Numerical Analysis, Theory of Computation.
- **Bachelor of Technology in Electronics and Telecommunication** Aug 2016—May 2020  
**NIT Raipur** – National Institute of Technology Raipur, Raipur, India. **GPA – 3.54/4.0**  
**Coursework:** Data Structures, OOPS(C++), Computer Networks, Computer Architecture, Neural Networks.

## TECHNICAL SKILLS

- **Data Analysis:** Exploratory Data Analysis (EDA), Data Cleaning, Feature Engineering
- **Tools:** Python (NumPy, Pandas, Matplotlib, Seaborn), SQL, Jupyter Notebooks
- **Statistical Analysis:** Hypothesis Testing, Regression Analysis, Time Series Analysis
- **Data Visualization:** Tableau, Excel
- **Machine Learning:** Scikit-Learn, TensorFlow, XGBoost
- **Deep Learning:** Neural Networks, CNN, RNN
- **Natural Language processing:** Experienced in NLP tasks, such as text classification, sentiment analysis, and named entity recognition
- **Database Management:** SQL, MySQL
- **Programming Languages:** Python
- **Version Control:** Git, GitHub

## CERTIFICATIONS

- IBM certified DATA SCIENTIST
- IBM certified DATA ANALYST

## WORK EXPERIENCE

### DATA SCIENTIST AT WIPRO LIMITED

Mar 2022– Mar 2023

- Successfully implemented a recommendation system in a production environment, which intelligently suggests additional items to customers based on their previous purchase history. This implementation resulted in a notable 17% increase in the average order value.
- Developed a Random Forest model for customer attrition that led to a 0.50% increase in monthly retention.

### DATA ANALYST AT WIPRO LIMITED

Sep 2020– Mar 2022

- Analyzed and interpreted large datasets to identify trends and insights that drive business decisions.
- Conducted exploratory data analysis on large datasets to identify trends, patterns, and insights.
- Managed data quality and data governance processes, ensuring data accuracy and completeness.
- Collaborated with cross-functional teams to define business objectives and data requirements.
- Managed projects related to data analysis and reporting, ensured timely and accurate delivery of results.

## ACADEMIC PROJECTS

### Space X Falcon 9 First Stage Landing Prediction

Jan2023 - Sep 2023

- As a Data Scientist, in this project I found the factors for high success rate in rocket landing. Collected data using SpaceX API and web scraping techniques.
- Performed Data cleaning and EDA. Explored complete data by visualization techniques. Analyzed data using SQL. Developed machine learning models to predict landing outcomes. [GitHub Link](#)

### Letter Dataset Analysis

- [GitHub Link](#)

### Boston Housing Dataset analysis:

- Analyzed correlations between key features and median house prices, identifying significant relationships. Utilized data visualization techniques to visually represent the impact of features on housing prices, enhancing clarity for non-technical stakeholders.
- Applied regression analysis to build predictive models for housing prices based on feature variables, achieving an R-squared value of 0.71. [GitHub Link](#)

### Decay-coefficients-in-a-Naval-propulsion-plant

- [GitHub Link](#)

### Dashboards using Excel:

- [GitHub Link](#)

### Various Machine Learning Models:

- [GitHub Link](#)

# Mohammad Beheshti

Address: 111B, N Stadium BLVD, APT 61, 65203

Mobile: (573) 810-1391

Email: mbv1377@gmail.com

---

## EDUCATION

- **Ph.D. in Health Informatics (With Emphasis on Data Science)** | May 2026 | **GPA: 4**  
University of Missouri, Columbia, MO, USA
  - **Master of Science in Health Informatics** | Dec 2023 | **GPA: 4**  
University of Missouri, Columbia, MO, USA
  - **Bachelor of Science in Health Information Technology** | Oct 2020 | **GPA: 3.94**  
Tabriz University of Medical Sciences, Tabriz, Iran
- 

## PROFESSIONAL EXPERIENCE

### Data Scientist (Research Assistant) | Missouri Cancer Registry and Research Center | Aug 2022 – Present

- Collaborate closely with the database team to perform the ETL processes on the cancer database and ensure accurate and timely data submission, enhancing the data pipelines and overall data quality.
- Conduct extensive statistical analysis using SPSS/SAS to identify cancer patterns in Missouri, identifying key trends and insights for further research and intervention.
- Employ supervised and unsupervised machine learning techniques to predict the effects of surgery on cancer patients, providing valuable insights for personalized treatment planning (using scikit-learn and TensorFlow).
- Implement advanced probabilistic linkage techniques on pathology data, increasing data completeness by 30% and securing the NAACCR Gold Certification for the cancer registry.
- Automate cancer reports and data preprocessing using Python, saving 80 hours of manual effort per month.
- Demonstrate proficiency in writing and executing complex SQL queries on a daily basis, facilitating data retrieval and analysis, and ensuring timely access to critical information for decision-making purposes.

### Population Health Analytic Intern | Centene Corporation | May 2023 – Aug 2023

- Conducted in-depth analyses examining the correlation between chronic conditions and healthcare costs, revealing key insights into cost drivers.
  - Performed geospatial analysis by accurately geocoding patient addresses to their respective census tracts.
  - Utilized Microsoft Power BI to create visually compelling dashboards that effectively communicated complex population-level data, enabling stakeholders to make informed decisions.
  - Prepared comprehensive reports and dynamic presentations summarizing findings, recommendations, and cost-saving opportunities, effectively communicating complex information to stakeholders.
- 

## SKILLS

- |                                                  |                           |                    |
|--------------------------------------------------|---------------------------|--------------------|
| • Python Programming                             | • R Programming           | • SAS and SPSS     |
| • Machine Learning (Supervised And Unsupervised) | • Microsoft Power BI      | • Tableau          |
| • SQL                                            | • Social Network Analysis | • Microsoft Office |
- 

## HONORS

- Student Employee Shout-Out, University of Missouri, 2022.
- The **2<sup>nd</sup> Best** General Magazine at the Ministry of Health Festival, 2019 (as the Editor-in-Chief).
- Dean's List (All Semesters).

# Adel Nasser Abdalla

(510) 9265576 | ana2169@columbia.edu | <https://github.com/aa221>

## EDUCATION

### Columbia University

*Masters in Data Science*

Coursework: Algorithms for Data Science, Computer Systems, Applied Deep Learning, NLP, Computer Vision.

New York City, New York

Expected Dec 2024

### University of California, Berkeley

*Bachelors in Data Science and Economics*

Coursework: Computational Structures in Data Science, Data Structure and Programming Methodology, Principles and Techniques of Data Science, Data Mining and Analytics, Statistical Theory and Methodology, Linear Algebra & Differential Equations, Introduction to Database Systems, Data Inference and Decisions, Experimental Design For Machine Learning.

Berkeley, CA

May 2023

## WORK EXPERIENCE

### UnitedMasters

*Data Engineer Intern*

San Francisco, California

Jun 2023 - Aug 2023

- Designed ETL pipelines using Python and DBT to migrate TikTok data from Snowflake into the company's Postgres Database.
- Automated the scheduling of said pipelines using Airflow to populate the Postgres database with updated TikTok data on a daily basis.
- Developed and routed APIs to access TikTok data on the company's track discovery page; boosting the A&Rs' ability to discover new, trending artists by almost 100%.

### Meta (Facebook)

*Data Science Intern*

New York City, New York

May 2022 - Aug 2022

- Improved video segmentation on the Facebook app by 3% (measured by co-consumption) by leveraging features of different video embedding methodologies and tuning hyperparameters of the current video segmentation model.
- Experimented with several unsupervised Machine Learning models to maximize video segmentation performance and achieve the highest possible similarity between videos in a given cluster.
- Built data pipelines to expand coverage of data available, and automate query operations for daily dashboards.

### Blitz Checkout

*COO and Co Founder*

Berkeley, California

Feb 2021 - Feb 2022

- Devised marketing and promotional tactics to grow the user base to 100+ customers.
- Designed and implemented a neural network to predict what customers will buy next, ensuring at least 70% of users were recipients of targeted promotions.
- Managed and directed two design interns hired through the Pad-13 Berkeley Skydeck program.

### The Economist Intelligence Unit

*Student Data Science Researcher*

Berkeley, California

Sep 2021 - Dec 2021

- Automated product onboarding processes by building web scrapers to gather product information from various retailers across 4+ countries and 10+ cities, and write said information into a database.
- Leveraged the Google Sheets API to translate scraped data into worksheet files for multiple cross functional teams to easily leverage said data.
- Expedited the entire product onboarding process by 200% and cut manual labor costs by a factor of 10.

## SKILLS

Programming Languages and Database Management.

- Python, Java, R, SQL, Tableau, DBT, Snowflake, Postgres, S3 (AWS), Pinecone (vector database).

Useful tools.

- Pandas, Selenium, Beautiful Soup 4, Numpy, Scrapy, Apache Airflow, Matplotlib, Boto3, EC2 (AWS), Flask, Github.

Machine Learning Models and Frameworks.

- Pytorch, Keras, Tensorflow, NLP Libraries, Scikit Learn, Leveraging Large Language models including BERT.

## PERSONAL PROJECTS

### Matching images to a songs based off of emotion—Image2Song

Sep 2023 - Present

- Utilized a LLM (LLAVA) to create a system capable of matching a given image to a song based on the emotion of both entities.
- Generated a data-set of over 8k songs and respective emotions and stored these as vectors in a vector database.
- Fine-tuned LLAVA to allow it to optimally perform for the image-song matching task, boosted model performance by 100% for the task.

### Album Similarity Visualizer — Album2Vec

Jan 2023 - Mar 2023

- Developed a visual representation of album similarities, where closer albums are more similar to one another.
- Generated a Vector Database of 25k+ albums and extracted the audio and musical features of each album by utilizing various music to feature APIs.
- Created a Neural Network to determine the weights of these 100+ musical features, optimizing for the similarity between clustered album genre.

**AMAN BHUTANI**  
**Boston, MA | +1 (857)296-2858**

Contact: [bhutani.am@northeastern.edu](mailto:bhutani.am@northeastern.edu) ; +1(857) 296-2858 LinkedIn: [linkedin.com/Ambhutani](https://linkedin.com/Ambhutani) GitHub: [github.com/Ambhutani](https://github.com/Ambhutani)

## **WORK EXPERIENCE**

<b>Data Science &amp; Artificial Intelligence Co-Op   Moderna Therapeutics   Cambridge, MA</b>	<b>July 2023 – Present</b>
- Developing Bayesian statistical models using Poisson-Gamma distributions to forecast patient enrollment for trial sites within 95% CI.	
- Implemented Bootstrapping, GLM, and Bayesian Neural Networks to generate generic site activation schedules for clinical trials.	
- Deployed predictive models using CI/CD through Bitbucket to be used by stakeholders across 9 departments at Moderna.	
- Conducted simulations for the models created to forecast total-time for completing clinical trials.	
<b>Graduate Teaching Assistant - Algorithms   Khoury College of Computer Science   Boston, MA</b>	<b>Jan 2023 – May 2023</b>
- Evaluated assignments and hosted weekly office hours for ~120 students to resolve their doubts about assignments and lectures.	
- Assisted professor in planning and implementing the program and daily curriculum activities of the course.	
<b>Data Science Intern - US Brokerage Operations   FlyHomes   Noida, India</b>	<b>June 2022 - Aug 2022</b>
- Performed extensive data pre-processing and wrangling in python to draw insights and prepare the company's data for modeling.	
- Improved home valuation tools by predicting correction index of Santa Clara County with 93% accuracy using Neural Networks & RFC.	
- Warehoused housing data in star schema and created weekly dashboards to track the housing market dynamics using Microsoft Power BI.	
- Created excel macros to facilitate analysis of a listing agent's past sales & geographical data helping sales teams send better applications.	
<b>Machine Learning Intern   August Infotech   Surat, India</b>	<b>Jan 2022 - May 2022</b>
- Utilized BeautifulSoup and Scrapy in python to create a dataset of hotel details, room & price information from 9 different websites.	
- Developed a Hybrid recommender system using content & location-based filtering to provide tailored options to the users.	
- Implemented location-based filtering using the Haversine formula to identify the top 15 hotels closest to the specified address & city.	
- Utilized cosine similarity to provide personalized recommendations based on user-specified preferences and hotel attributes.	

## **PROJECTS & HACKATHONS**

<b>Predicting S&amp;P500 movement using Time-Series Analysis &amp; Machine Learning [<a href="#">Link</a>]</b>	<b>Jan 2023 – May 2023</b>
- Pre-processed data acquired from YFinance including opening & closing prices and calculated various market indexes in Python.	
- Implemented TimeSeriesForest, XGBoost, & SVM to predict the direction and magnitude range of the movement of S&P 500.	
- Devised experiments to test performance of model predictions over various time frames and change in magnitude of S&P 500 index.	
<b>Data mining &amp; analysis for 10 years of journal publication data - DBMS (CS5200) [<a href="#">Link</a>]</b>	<b>Oct 2022 - Dec 2022</b>
- Designed a normalized relational schema and created an ETL pipeline to import data from XML file into the relational database in R.	
- Utilized AWS - RDS service to create a star schema and populated the warehouse using the data stored in the relational database.	
- Queried data from the warehouse using SQL commands to conduct analysis and visualize results obtained using ggplot in R.	
<b>Effects of Dropouts on Neural Networks [<a href="#">Link</a>]</b>	<b>May 2021 - Aug 2021</b>
- Studied effects of changing dropout rates on performance of Artificial & Convolutional Neural Networks in Image and Text classification.	
- Designed and executed A/B experiments to study impact of dropouts on various hyperparameters & success metrics.	
- Utilized LIME to create surrogate models to study the relation of changing dropout rates on explainability of black-box models.	
<b>COMPAS dataset bias analysis - Fairness in Machine Learning [<a href="#">Link</a>]</b>	<b>Jan 2021 - Apr 2021</b>
- Pre-processed and analyzed COMPAS dataset used for rating an individual for their likelihood of recidivating using Pandas and NumPy.	
- Leveraged various bias prevention techniques to create efficient risk models using SVM, RFC, and Artificial Neural Networks in Scikit.	
- Reduced false positive rate of prediction by 11% and found features most contributing to bias against minority groups in the dataset.	

## **TECHNICAL SKILLS**

<b>Language &amp; Libraries</b> - Python, R, C/C++, MATLAB, Scikit-learn, TensorFlow, Keras, PyTorch, Pandas, NumPy, PyMC
<b>Database &amp; Other Tools</b> - MySQL, Docker, Bitbucket, Git, AWS, MS Excel, PowerBI, Tableau

## **EDUCATION**

<b>Northeastern University - Khoury College of Computer Sciences, Boston, MA</b>	<b>Sep 2022 - Dec 2024 (Anticipated)</b>
MSc in Artificial Intelligence - 3.7/4.0	
Relevant Coursework - Machine Learning, Statistical Inference for Data Analysts & Engineers, DBMS, Algorithms	
<b>Navrachana University, Vadodara, India</b>	<b>May 2018 - May 2022</b>
Bachelor of Technology in Computer Science & Engineering - 8.7/10	

**ELISHA (ELI) EDME**  
(516)-330-8657 | [eae8374@nyu.edu](mailto:eae8374@nyu.edu) | [LinkedIn](#) | [Github](#)

## EDUCATION

---

**New York University, Tandon School Of Engineering**, New York, NY May 2025  
*Bachelor of Science in Computer Science, Minor in Finance and Mathematics (GPA: 3.4)*  
Relevant Coursework: Machine Learning, Projects in Data Science, Applied Cryptography, Computer Architecture, Databases, Linear Algebra, Data Analysis, Financial Statements, Corporate Markets, Object Oriented Programming

## PROJECTS

---

- Student Course Scheduler** | *JavaScript, React, MongoDB, jQuery* Sep. 2023 – Present
- Constructed application to find and rank possible student schedule given student preferences
  - Automated class addition to MongoDB database by web scraping from school public records website
  - Experimented with 7 ranking algorithms to boost schedule preference accuracy by 56%
- NFL Wins Predictor** | *NumPy, Sklearn, Pandas, Matplotlib* Jan. 2023 – Jan. 2023
- Generated feature vector using Pandas Dataframe to store relevant features of NFL statistics
  - Visualized NFL statistics with Matplotlib and Seaborn to accurately tune hyperparameters
  - Leveraged Linear Regression model to build a Machine Learning algorithm with 90% accuracy
- Enigma Machine Simulator** | *Python, Flask, Bootstrap HTML/CSS, Figma* Sep. 2022 – Dec. 2022
- Developed a digital replica of a encryption machine to visually represent device's functionality
  - Leveraged modular arithmetic to reduce app runtime by 23 ms without compromising accuracy (97%)
  - Utilized Bootstrap HTML/CSS to build flexible, user-friendly interface accessible on any device

## EXPERIENCE

---

- Co-Founder, Cyber (Startup Company)**, Brooklyn, NY Jul. 2021 – Mar. 2022
- Formed team of 5 students to create a service giving users critical medical advice with Machine Learning
  - Researched popular Webflow designs and constructed cohesive typescript form to acquire user symptoms
  - Analyzed previous Machine Learning code to make decision tree model for diagnosis, with 83% accuracy
  - Studied how to lead, market, and plan during entrepreneur workshops; improved productivity by 45%
  - Achieved 2nd place in Entrepreneurship Competition challenge hosted by NYU Tandon
- Member, NYU Augmented Library Project Team**, Brooklyn, NY Sep. 2021 – Dec. 2021
- Collaborated with 2 students to develop a 10 question user-survey form to enhance app performance
  - Experimented with 2 Virtual Reality libraries to use for guiding students through school library
  - Categorized and formatted user satisfaction responses for easy readability; increased productivity by 16%
- Software Intern, KJ Technologies**, New York, NY Jun. 2021 – Aug. 2021
- Partnered with advisor to create python scripts to maximize in-office efficiency
  - Analyzed 20+ custom requests from customers while setting up office devices
  - Collaborated with different departments to gain a comprehensive understanding of software/IT industry
  - Updated Microsoft Excel spreadsheets to accurately record current inventory of 150+ items

## TECHNICAL SKILLS

---

Programming Languages: C/C++, Python, Java, JavaScript, HTML/CSS, Typescript  
Frameworks/Libraries: React, MongoDB, Flask, jQuery, NumPy, Pandas  
Developer Tools: Github, VS Code, Visual Studio, PyCharm, IntelliJ, Eclipse

# BHAVIK AGARWAL

Maryland, United States

📞 +1 (667) 899-1927 📩 [bhavik2506@gmail.com](mailto:bhavik2506@gmail.com) 💬 [linkedin.com/in/bhavik-agarwal-7a5432178/](https://linkedin.com/in/bhavik-agarwal-7a5432178/)

## Education

---

<b>Johns Hopkins University</b> <i>Masters of Science in Engineering in Computer Science</i>	<b>August 2023 - Present</b> CGPA: 4.0 / 4.0
<b>IIIT Delhi</b> <i>Bachelor of Technology in Computer Science Engineering and Biosciences</i>	<b>August 2018 - April 2022</b> CGPA: 3.8 / 4.0
<b>Sachdeva Public School, Rohini</b> <i>Higher Senior Secondary (CBSE)</i>	<b>April 2017 - April 2018</b> Percentage: 93.2%
<b>Sachdeva Public School, Rohini</b> <i>Senior Secondary (CBSE)</i>	<b>April 2015 - April 2016</b> CGPA: 10.0 / 10.0

## Experience

---

<b>Google Research</b> <i>AI/ML Scholar</i>	<b>Sep 2023 - Present</b> California, United States
• Selected by Google in a class of 40 students out of over 50k applications to be mentored by people at Google Research to advance my career as an ML/AI/Software Engineer and conduct novel research during the program.	
<b>T.I.R.I. Lab - JHU (Project Funded by Microsoft Research)</b> <i>Graduate Research Assistant - NLP</i>	<b>Aug 2023 - Present</b> Maryland, United States
• Developing a Multimodal Healthcare Large Language Model (Natural Language Processing, Deep Learning) over various kinds of medical records, literature utilizing the GPU architecture of Johns Hopkins University.	
<b>BlueStacks - now.gg</b> <i>Member of Technical Staff (MTS) - AI/ML</i>	<b>July 2022 - July 2023</b> Delhi, India
• Deployed Recommendation system for games and ads using collaborative filtering, content filtering, and hybrid methods (matrix factorization and neural network). Achieved more than 95 % accuracy during user analysis on BlueStacks App Player. • Created live anomaly alert tracking for traffic during various funnel events using RNN on a user base of millions over various countries.	
<b>BlueStacks - now.gg</b> <i>SWE (AI/ML) Intern</i>	<b>Jan 2022 - July 2022</b> Delhi, India
• Deployed Deep Learning Models such as LSTM and N Beats for univariate and multivariate time forecasting of large varied amount of user traffic throughout the world. Cutting down the cost paid by the company i.e. millions of dollars by predicting beforehand the optimal number of servers to run so that unserved and idle count is reduced. Managed more than 99% availability and less than 1% idle till now which is considered a state-of-the-art benchmark.	
<b>Western Digital - SanDisk</b> <i>Engineering Intern</i>	<b>June 2021 - August 2021</b> Bangalore , India
• Programmed various unit test cases in backend architecture for reading, updating and removing the data given by the host system in the NAND and NOR chips (Single Structure Architecture, Multiple Structure Architecture) through SEM Queue using Hashing in Storage Management Tables. The main purpose was to utilize the RAM as min as possible.	
<b>Amdocs</b> <i>SWE Intern</i>	<b>May 2021 - June 2021</b> Remote Work
• Worked with Amdocs Global Team (USA, Israel, India). Automated the analysis of manager-employee bidirectional feedback using a data pipeline and reduced the time to analyze by 40%, a significant improvement to the then existing MS-Excel based solution.	

## Publications

---

**Exploration of Negative Mass as a Harbinger of Dark Energy.** IOSR Journals. | [Paper](#) | *Agarwal B. and Babbar A.*

**Time Series Nowcasting of India's GDP with Machine Learning.** IEEE Conference. | [Paper](#) | *Agarwal B. and Malik N.*

**Plant-pollinator network analysis.** RIISE 2022 | [Poster](#) | *Agarwal B., Gupta R. and Dr. Swapna Purandre*

**Sarcasm Detection in News Headlines using NLP** | Paper - submitted | *Agarwal B., Chawla N. and Dr. Narina Thakur*

**Clickbait Detection using hybrid BERT-CNN-LSTM** | Paper - in progress | *Agarwal B., Chawla N. and Dr. Chavi Dhiman*

## Projects

---

**Doping Analysis** | *Python, PyTorch, SciKit, Jupyter* | [Presentation](#) | *Team Size:6*

- The objective was to identify cases of unintentional doping of sportspersons. Achieved 94.9% accuracy through the random forest algorithm in predicting the likelihood of a particular protein being targeted by a banned drug in the future through drug-protein pathway analysis. Pre-existing standard datasets did not exist. We scraped the web and analyzed banned drugs (i.e., their properties and pathways) to create a database of all the known drugs for analysis. Identified alternatives to banned drugs and compounds using combinations of properties. Work revealed that many top athletes were wrongly banned from sports due to doping and we have sent our report to higher authorities for investigation.

**Food DB - Resource for carbon and water footprint of recipies** | *React Native, Google Firebase* | [Website](#) | *Team Size:3*

- Website created to find out the carbon footprint and water footprint of various recipes throughout the world. All the workflow was created in python and we managed to increase the speed and efficiency of previous implementation by using hash maps and optimizing library usage – total runtime reduced from 571s to 121s. Also created the frontend of landing page and integrated the frontend to backend.

## Honours and Achievements

---

- Google's Prestigious CS Research Mentorship Program acceptance out of many candidates
- Dean's List of Academic Excellence awarded by Dean of Academics, IIIT Delhi
- Dean's List of Academic Excellence for excellent growth in academics awarded by Dean of Academics, IIIT Delhi
- Solved 1000+ coding problems on various platforms like Leetcode, Codeforces, Codechef, etc.
- Top 1% of 1.5 million aspirants in the Indian Joint Entrance Exam (JEE)
- Secured 2nd position at Inter Zonal held at National Level for Table Tennis.
- Awarded best student in high school for all-round performance.

## Technical Skills

---

**Languages:** Java, Python, C++, C, JavaScript

**Technologies/Frameworks/Libraries:** PyTorch, Tensorflow, Keras, GCP, React JS, Node JS, Selenium, JavaFX, MySQL, Django, Flutter, Google Firebase, Mongo

**Technical Electives:** Natural Language Processing, Data Structures And Algorithms, Operating Systems, Database Management Systems, Statistical Machine Learning, Computer Networks, Big Data Mining in Healthcare, Advanced Data Structures and Algorithms, Algorithms for Bioinformatics, Blockchain and Cryptocurrency

## Leadership and Extra Curriculars

---

- Student Mentor (Mentoring Fresher Students for their academic and day to day life challenges)
- Teaching Assistant for Statistical Machine Learning Course (Conducted tutorials, created assignments, and exams for 100+ students on various statistical machine learning topics such as Random Forest, K Means, Deep Neural Networks, etc.)
- Teaching Assistant for Algorithms in BioPhysics Course (Conducted tutorials, created assignments, and exams for 100+ students on various algorithmic topics in the field of bioinformatics like BLAST, Needlman-Wunsch etc.)
- Coding Instructor, IIIT Delhi (Data Structures Problem Tutorials uploaded online)
- Sports Coordinator, Table Tennis, IIIT Delhi
- Head Prefect and Sports Captain , Sachdeva Public School

---

Alessandro Hizdri

Birthday: 06/06/1998

Birth place: Alessandria/ Italy

1925 E Sergeant Street • Philadelphia, PA 19125 • alehizdri3@gmail.com • +39 3891978858

## EDUCATION

### POLYTECH INC OF TURIN

Fall 2023

- Master of Science, Biomedical Engineering and E-health.
- Selected for Erasmus + program, in Paris, France.

### POLYTECH INC OF TURIN

March 2021

- Bachelor of Science, Biomedical Engineering
- Invited to: Maker Faire Rome-European edition (Fall 2019)
- Final Project: Fixation devices for traumatology Surfaces' focus: State of art and possible future solutions Overall Grade 97 /110 (3.3 GPA)

## EXPERIENCE

### ENGINEER – INFOTAINMENT ENGINEER FOR PORSCHE

- March 2023 Consultant at Italdesign, Via Achille Grandi 25, 10024 Moncalieri, Turin, Italy.

Automotive Engineer specializing in **Test and Validation**, aiming to ensure optimal performance and safety of vehicles through a comprehensive range of testing procedures.

- Daily setup of Porsche benches and vehicles setup with weekly report of conditions of the benches.
- Software upgrade and evaluation with punctual reporting to customer and suppliers
- Hardware upgrade and installation to improve cars/benches setup
- Coordination between department to assure proper testing of new software releases
- Creation, execution and automatization of smoke tests, stability test and performance test in a weekly cadence to respect deadlines for software releases, preparation of **documentation and release emails** to be addressed to the suppliers in order to fix problems on past releases.
- Log acquisition via Esotracer, teraterm, foxterm and similar from different ECU of cars/benches
- Defect reporting, creation and follow-up through KPM , Jira
- Daily interactions with different teams for coordination of products delivery.
- Business trip for drive testing and workshops at Porsche headquarter.

### BIOMEDICAL ENGINEER - IMAGING INTERN

- 2021/2022 Télécom ParisTech, 19 Pl. Marguerite Perey, 91120 Palaiseau, France.

During my internship, I embarked on a cutting-edge research project centered around knowledge distillation applied to medical imaging, specifically in brain tumor segmentation.

- **Literature Research and Method Exploration:** in-depth literature review to understand the prevailing trends and challenges in the domain of medical image segmentation.
- **Adapting Attention Transfer and Contrastive Learning:** exploration and adaptation of novel techniques to enhance the performance of the initial brain tumor segmentation network. investigating attention transfer methods and contrastive learning techniques, which had not been previously tested in this context. This process involved understanding the underlying principles, reviewing related research papers, and adapting these techniques to suit the medical imaging domain.
- **Deep Network Hyperparameter Tuning:** selection of attention transfer and contrastive learning approaches, fine-tuning of the **deep neural network's hyperparameters**. For the optimization of the performance of the "student" network while effectively leveraging the knowledge from the "teacher" network. Tools such as **TensorFlow and scikit-learn** have been explored, iteratively adjusting hyperparameters to strike the right balance between accuracy and training speed.

- **Utilizing Libraries and Tools:** all the project have been developed using **python** and working on **Linux environment** for the tuning of the framework through the use of the Company's clusters. Extensively employed libraries such as **TensorFlow** for building and training neural networks, **scikit-learn** for hyperparameter tuning, and **Pandas** for data manipulation and analysis. TensorFlow's ecosystem provided the necessary tools for constructing and training the intricate network architecture, while scikit-learn streamlined the hyperparameter optimization process. Pandas enabled efficient handling and analysis of the medical imaging dataset, further enhancing the project's efficiency.
- 

## PROJECTS

### ANDRACING

- 2019 Hackability, No-profit organization, Via Bezzecca, 11, 10131 Torino TO

As part of my extracurricular activities, I worked with a non-profit start-up called Hackability, with the aim of designing a customized kart for children with disabilities. My team and I **leveraged 3D printing technology** to create pedals, support elements, and a new seat tailored to the specific needs of the patient. We were extremely **proud to present** our project at the **Rome Maker Faire**, one of the most important Italian stages for start-ups, students and researchers with an **international visibility**, in the fall of 2019.

### DIGITAL PATHOLOGY CHALLANGE

As part of a digital pathology challenge project, I worked on **developing deep learning techniques** to segment single-cell nuclei in hematoxylineosin stained images of colon sections. The course demanded the implementation on **Google Colab** enviroment of a **convolutional neural network pre-trained** with various methods of pre and post-processing (**data augmentation, Watershed segmentation**) to address the challenges arising from unbalanced, diverse datasets and to correct mistakes made by our **segmentation model**. Through this project, I was able to gain valuable experience in working with **large datasets**, as well as **advanced image segmentation** techniques and further develop my **programming skills in python**.

### DATA SCIENCE IN MEDICINE

As part of a group project, I tackled the challenge of **developing computer-aided diagnosis (CAD) and computer-aided detection (CADx)** tools for **the segmentation and prediction** of aggressiveness of prostate cancer in t2-weighted magnetic resonance images and diffusion maps. The project has been developed on **MATLAB** and involved a population of 60 patients, allowed the collection of data and the use of machine learning methods for **feature extraction, feature selection, and classification tasks** were used in order to build the models above mentioned. I **evaluated the performance** of different classification methods and gained a deeper understanding of how to optimize machine learning models for medical imaging applications. Through this project, I was able to hone my skills in **data analysis**, feature engineering, and machine learning model selection.

### CONTACT TEMPERATURE GAUGE

Design of an electronic system aimed at measuring contact temperature, utilizing **Arduino R1 and Amtel 7** for coding. The validation process involved testing the system's accuracy, precision, and reliability under various conditions using **specialized instruments and techniques**. The project provided an opportunity to develop skills in electronic circuit design, coding, laboratory testing, project management, and problemsolving.

## SKILLS & INTERESTS

**Programming languages:** C++, MATLAB, Python.

**Equipment:** multimeters, analogue and digital, oscilloscopes, and under some circumstances more specialized test equipment including spectrum analyzers.

**Tools:** Confluence, KPM, Jira, PIDT, Esotraceviewer, TeraTerm, Confluence, KPM, Jira.

**Data analysis and visualization:** Pandas, NumPy, Matplotlib, Excel.

**Machine learning frameworks:** Scikit-learn, TensorFlow, Keras, PyTorch.

**Machine learning algorithms:** Classification algorithms, Regression algorithms, Clustering algorithms, Dimensionality reduction algorithms, Neural Networks, Bayesian learning algorithms.

**Language:** Albanian (Mother Language), Italian (Mother Language), English (Fluent), French (beginner).

**Communication and teamwork:** Communication, Teamwork, Problem-Solving, Time Management, Leadership, Creativity, Critical thinking, Conflict resolution, Adaptability, Determination.

# Marielys Reyes Castillo

mr258@njit.edu | (908) 328-6183 | Elizabeth, NJ | [www.linkedin.com/in/mrc258](https://www.linkedin.com/in/mrc258)

## EDUCATION

New Jersey Institute of Technology  
B.S. in Biomedical Engineering  
Minor in Applied Mathematics  
Current (Cumulative) GPA: 3.0

Expected Year of Graduation: May 2025

## AFFILIATIONS

Society of Hispanic Professional Engineers Event Coordinator Fall 2023

## RELEVANT COURSEWORK

Fundamentals of Engineering Design| Differential Equations| Calculus | Biological and Chemical Foundations | Mechanical Fundamentals | Material Fundamentals| Probability and Statistics| Electrical Fundamentals

## KEY SKILLS

Microsoft Efficient| MATLAB Certified| Communication |Critical Thinking| C++| Bilingual |GrabCAD | Cura

## EXPERIENCE

Department of Biomedical Engineering -NJIT Sep 2022-Present  
Lab Technician Newark, NJ  
- Examined job orders to determine quality, colors, and printing instructions  
- Created estimates for parts and materials, set up machines, and scheduled prints

## LEADERSHIP EXPERIENCE

Educational Opportunity Program-NJIT Sep 2021-Present  
Office Coordinator Newark, NJ  
- Supervised operations to increase the distribution of information to a multitude of people with efficiency  
- Managed teams to reach deadlines by delegating tasks to the best-suited team member  
- Screened and transferred incoming calls and transmitted information and documents to internal personnel  
- Managed interviews and recorded and typed all responses of the interviewee

## RESEARCH

Garden- State Louis Stokes Alliance for Minority Participation (GS-LSAMP) Summer 2023  
Department of Biomedical Engineering-NJIT  
- Used neuroimaging techniques in MATLAB to calculate the amplitude of low-frequency fluctuations and the regional homogeneity of subjects who are on the autistic spectrum compared to subjects who were not

## ACADEMIC PROJECT

Robotic Hand- NJIT Fall 2021  
- Design Goal: Create a 3D-printed hand for a user who needs one  
- The 3D-printed hand would mimic the motions of the user's functioning hand with the help of servo motors and flex sensors sewn onto a glove  
- The user would attach the glove to their functioning hand so the 3D-printed arm can replace the non-functioning one, which won in the First-Year Design Showcase in February 2022

## VOLUNTEER EXPERIENCE

STEM NOLA Event | Bronx, NY May 6, 2023  
- In-person STEM Fest hosted by STEM NOLA and 21 Cent Ed, introduced 500+ K-12 students  
Dia De Ciencias | Newark, NJ March 25, 2023  
- Hosted workshops and taught K-12 students about STEM

# Sean Neuman

sneuman157@gmail.com | (484)-838-9118

605 Ridge Rd Spring City PA, 19475

## EDUCATION

**Pennsylvania State University** | State College, Pennsylvania

Expected Graduation: May 2025

*Bachelor of Engineering:* Computer Science

GPA: 3.0/4.0

**Malvern Preparatory School** | Malvern, Pennsylvania

Graduated: May 2021

GPA: 3.80/4.0

## NOTABLE COURSEWORK

### Artificial Intelligence

- ❖ Apply basic principles of AI in solutions that require problem solving, inference, perception, knowledge representation, and learning.
- ❖ Demonstrate awareness and a fundamental understanding of various applications of AI techniques in intelligent agents, expert systems, artificial neural networks and other machine learning models.

### Data Structures and Algorithms

- ❖ Determined the growth rates of functions and analyzed the complexity of various algorithms.
- ❖ Solved recurrence relations when calculating time and space complexity.
- ❖ Utilized a variety of algorithms and design paradigms.

### Object-Oriented Programming with Web-Based Applications

- ❖ Introduced virtual machines, programming language concepts and web-based programming using Java.
- ❖ Object-oriented design, event-handling methods, Web technologies, virtual machines, graphical user interfaces, API programming.
- ❖ Created a course scheduler app that managed both student and admin requests using a Java GUI and utilized SQL calls to a database.

## WORK EXPERIENCE

Penn State Facilities, *Esports Attendant*

October 2023 - Present

- ❖ Responsible for overseeing the daily operations of the computer lab.
- ❖ Diagnostics and Maintenance of over 40 high-performance computers dedicated to Esports.

Camp Rainbow, *Head Lifeguard and Aquatics Director*

June 2019 - August 2019

- ❖ Lifeguarding, maintained pool and facilities.

Gentle Dentist, *Dental Assistant*

March 2020

- ❖ Worked with 3D printers to create molded dental aligners (Invisalign) and IT work.

## ADDITIONAL EXPERIENCE

Malvern Prep Engineering Club, *Team President, Head Mechanic*

2017 - 2021

- ❖ Lead FIRST Robotics Competition, where we placed 1<sup>st</sup> in our district event in 2018.
- ❖ Mentored the middle school engineering club on a weekly basis.

PSU Hackathon, *Project Leader*

November 2022 & April 2023

- ❖ Used an open-source eye tracking software to manipulate and simulate mouse movements and clicks on a computer designed for people with limited motor capabilities.
- ❖ Made an app that utilized AR to digitally paste photos onto a white t-shirt for design prototyping.

MetaCTF

January 2020, 2021, 2022

- ❖ An annual cybersecurity “Capture the Flag” competition where a competitor has 24 hours to solve a variety of challenges to find hidden pieces of data called “flags”, topics included web exploitation, forensics, cryptography, reverse engineering and more.

## SKILLS

Computer Skills: Fluent in Java, Python, C++, and C. Experience in MIPS, Verilog and SQL.

Advanced in Microsoft Office Suite, Linux, Photoshop, SolidWorks, and Excel

Intermediate in French

# Amogh Patankar

408-597-2478 | [apatankar@ucsd.edu](mailto:apatankar@ucsd.edu) | [linkedin.com/in/apatankar22](https://linkedin.com/in/apatankar22) | [github.com/apatankar22](https://github.com/apatankar22)

## EDUCATION

### University of California, San Diego

M.S., Computer Science & Engineering; Concentration: Artificial Intelligence & Machine Learning

Sept. 2023 – June 2025

La Jolla, CA

### University of California, San Diego

B.S., Data Science

Sept. 2020 – March 2023

La Jolla, CA

## EXPERIENCE

### Research Scholar

Stanford University School of Medicine

June 2023 – Present

Palo Alto, CA

- Carried out research under Dr. Eric Gross to observe and analyze opioid related overdose mortality of Asian-Americans and other races using the NVSS database.
- Conducted research under Dr. Latha Palaniappan and Dr. Jin Long to examine the effectiveness of strength training for diabetics amongst Asian-American subgroups in controlling hemoglobin levels (glycemic control).

### Software Development Engineer Intern

Amazon Web Services (AWS)

June 2022 – Sept. 2022

Seattle, WA

- Worked in AWS Lex, an organization that creates conversational chatbots powered by AI and Machine Learning.
- Recommended architectural changes for Lex ASR (Automatic Speech Recognition) Service and AWS DataHub.
- Designed and enhanced ASR and DataHub schemas to allow storage of non-critical data of customers in DataHub.
- Optimized ASR Service to allow faster resolution for customer requests by ~75%.

### Research Intern

Scripps Research Translational Institute

June 2021 – Aug. 2021

La Jolla, CA

- Developed an R library to estimate genetic regulatory variation using a confidence interval estimation method.
- Implemented various statistical concepts like binomial distributions and parametric bootstrapping, and applied them to data from the Genotype-Tissue Expression Project (GTEx).

### Firmware Embedded Engineering Intern

Inphi Corporation

July 2019 – Sept. 2019

Santa Clara, CA

- Developed Python modules and a user interface to create and display data based on test options and feature selection.
- Developed firmware in C++ to parse data, removing duplicates based on timestamp & hex value, per modular architecture.

### Senior Mathematics Tutor

De Anza College- Math, Science, and Resource Tech Center

Jan. 2019 – Aug. 2020

Cupertino, CA

- Led individual and group tutoring sessions in algebra, pre-calculus, calculus, and linear algebra.
- Mentored a group of tutors, and was heavily involved in center resource evaluation, and talent selection team.

## SKILLS

- Languages: Python, Java, R, C++, SQL, C#, JavaScript
- Frameworks: PyTorch, Tensorflow, Dask, PySpark, Keras, numpy, scikit-learn, pandas
- Misc: CI/CD, Docker, Software Development

## PROJECTS

- **Capstone Project:** Active Learning with Neural Processes for Epidemiology Modeling
- Autonomous Vehicle Trajectory using Deep Learning (**Argoverse 2 Dataset**, PyTorch)
- Data Science Interview Tool (GPT-3, Python)
- Predicting NBA Attempted Shot Success using Neural Networks (TensorFlow)
- Machine Learning Recommender System for Predicting Car Resale Price (scikit-learn, XGBoost, LightGBM)

## COURSES

CSE 151B, 158: Deep Learning, Web Mining and Recommender Systems

DSC 140A, 140B: Probabilistic Modeling and Machine Learning, Representation Learning

LIGN 167: Deep Learning for Natural Language Understanding

DSC 100, 102, 106: Data Management, Systems for Scalable Analytics, Data Visualization

DSC 180AB: Capstone Data Science Project

COGS 108: Data Science in Practice

MATH 183, 189: Statistical Methods, Exploratory Data Analysis and Inference

Sean O'Leary  
[soleary@stevens.edu](mailto:soleary@stevens.edu) | Hoboken, NJ | 732-425-2794

## Education

---

### **Stevens Institute of Technology | Hoboken, NJ**

Bachelor's Degree in Quantitative Finance | GPA 3.5

Minor in Computer Science | GPA 3.8

**Expected 2025**

**Relevant Coursework:** Stochastic Calculus, Advanced Probability, Statistics, Pricing and Hedging, Machine Learning and Advanced Time Series Analysis, Multivariate Calculus for Optimization, Linear Algebra, Data Structures and Algorithms, Corporate Finance, C++ for Finance

## Work Experience

---

### **Virtu Financial**

*Ondesk Engineering Intern | New York City, NY*

5/23-8/23

- Actively participated in quant training exercises and lectures, increasing my analytical skills and problem-solving abilities
- Delivered technical guidance and support to users across multiple locations, managing both virtual and physical desktop environments

### **Stevens Student Managed Investment Fund**

*Quantitative Strategist | Hoboken, NJ*

8/23-5/24

- Leading a small team through the research, development, and optimization of an enhanced pairs trading algorithm
- Implementing the OPTICS machine learning technique to identify and form robust trading pairs that maintain their relationship through volatile market conditions
- Utilizing XgBoost, a powerful machine learning algorithm, to amplify the predictive capabilities of our trading model, leading to more accurate trade execution and risk management.

### **Stevens Hanlon Financial Systems Laboratory**

*Research Assistant - SHIFT Project*

8/23-Present

- Working with the Automated Market Making research team to adapt the current SHIFT system to realistically simulate cryptocurrency exchanges
- Pioneering the creation and implementation of innovative mathematical models to optimize the AMM's fee structure, enhancing liquidity provision and overall efficiency

### **Stevens Institute of Technology**

*Computer Science Teaching Assistant | Hoboken, NJ*

1/23-5/23

- Helped teach students the core concepts of Computer Science and the Python language by hosting a weekly lecture.
- Graded assignments, debugged code, and developed lesson plans for a group of about 40 students.
- Worked with the professor and other assistants to ensure lesson plans were coordinated and impactful
- Contributed to the academic success of students, with students from my sections being more than three times as likely to receive an A grade than other sections

## Key Projects

---

### **Twitter Sentiment Pricing Model**

- Developed a pricing model using Twitter data as a proxy for market sentiment to predict a company's daily return using a Natural Language Processing model to score sentiment values
- The trading strategy developed using the model had roughly 3% excess returns above the market every year from 2010-2020

## Skills & Certifications

---

**Programming & Software:** Python, R, C++ (Beginner), Bloomberg Terminal, MS Office

# Zubair Lakhia

[lakhia@chapman.edu](mailto:lakhia@chapman.edu) • 408-913-0497 • <https://bit.ly/Zubairln> • <https://github.com/ZubairL>

## EDUCATION

### MSc, Data Science & Analytics

Georgia Tech, Atlanta, GA • Computational Analytics Track

Aug 2023- Present

### BSc, Molecular Biology, Double Minor in Data Science & Neuroscience

Chapman University, Orange, CA • Hesperian Scholar

Aug 2018- May 2022

## LABORATORY EXPERIENCE/ EMPLOYMENT HISTORY:

### Informatics Fellow (Contract) @ PTC Therapeutics | South Plainfield, New Jersey

June 2022- June 2023

- Demonstrated fluency in D360(data viz platform) and mentored scientists on querying/visualizing data
- Optimized a SAR Table pipeline with Microsoft VBA and Excel
- Created Python scripts for file splitting 96-Well data
- Assisted in data migration process between multiple LIMS/ELN informatics platforms

### Data Science Intern @ SciLifeLab | Stockholm, Sweden

Summer 2021

- Performed a comparative bioinformatics analysis of proteomics data to investigate relationships between exercise training, aging, and pathological conditions
- Refined and modified a Limma Analysis pipeline in R to measure and identify differentially expressed proteins

### Software/Biomedical Engineering Intern @ Keck Graduate Institute | Claremont, CA

Summer 2020

- Assembled an image processing/analysis pipeline with Python and OpenCV to observe and analyze micro-bead size, fluorescence, and color for multiplexed NA (nucleic acid detection)
- Developed pipeline was successfully able to predict and determine pathogen identity/concentration for UTI's
- Engineered and optimized a micro-bead mixing chamber geometry to selectively capture different nucleic acid sequences for Urinary Tract Infections (UTI's)

### Developmental Biological Research Intern @ San Francisco State University | San Francisco, CA

Summer 2019

- Investigated how the membrane bound Jagun protein affects the pathway in which the E.R. (Endoplasmic Reticulum) splits during mitosis in Drosophila (fruit flies) eyes

## TECHNICAL/LABORATORY TOOLS

Excel • Word • Docs • PowerPoint • Microsoft Suite • Python • R • SQL(PostGre)• Pandas • Matplotlib • NumPy • SciKit Learn  
• Plotly • Seaborn • Onshape(CAD) • Ansys • GENEious • Jupyter Notebooks • Machine Learning(ML) • LIMS/ELN(Dotmatics & Sapi) • Bioconductor • D360 • AWS(Lambda,S3,EC2,Cloudshell) • Bash

## PROJECTS:

### End to End Data Engineering Project-Spotify ETL Pipeline

July 2023

- Developed automated ETL pipeline with Spotify API, AWS (Lambda, Glue,S3,Athena,Quicksight)

### Evoked Action Potential Firing Rate Analysis with Python

Fall 2020

- Programmed logistic/ linear regression models using data from the Allen Brain Cell Institute database to compare the evoked action potentials of aspiny and spiny neurons in mice and humans with python

### Propensity Modeling- Rijpen

Apr/May 2021

- Managed a group of 3 students to create multiple propensity models
- Incorporated and built regression/tree-based models to predict the probability of companies being able to acquire other companies

### Scaled-Down Centralized Autonomy

Fall 2018-Spr 2020

- Built an RC car with lane and object detection with python
- Aim of project was to eliminate drunk, drugged and distracted driving

## PROFESSIONAL/VOLUNTEER EXPERIENCE:

### Data Science and Computational Biology Panelist @ SparkED ||Remote

Summer 2020

- Sharing my journey, experiences, and insights with target audiences of 30-40 high school students
- Offered one on one support with diverse students who are interested in data science and bioinformatics

### DS4A 2020 [Data Science For All Program Ambassador] @ Correlation One ||Remote

Oct 2020-Dec 2020

- Advertised and promoted curriculum that included exposure to full stack data fluency, data analytics, machine learning and AI

## AWARDS/CONFERENCES/MISCELLANEOUS

- Poster Presentation - Biomedical Engineering Society Meeting 2021 Orlando
- Co-author Publication– Transcriptomics/Proteomics Analysis of Exercise Training and Diseased Patients

# CHRISTINA CHANDRA

1334 Rosen Rd | Atlanta, GA 30319 | 404-219-1078 | [christina.chandra@gmail.com](mailto:christina.chandra@gmail.com) | [LinkedIn](#) | [GitHub](#)

## SUMMARY

As a fourth-year PhD candidate in Epidemiology with substantial training in data analysis and management, statistical analysis, and infectious disease modeling, I am excited to explore industry opportunities that allow me to transfer my programming and analytical skills to complex business problems. Given my prior breadth of business and academic research experience, I have experience communicating effectively with both broad and leadership audiences and experience collaborating and building relationships in cross-functional teams.

## EDUCATION

<b>Emory University</b>	<b>Atlanta, GA</b>
<i>Doctor of Philosophy (PhD), Epidemiology (GPA 4.00/4.00)</i>	Expected May 2025
• Relevant coursework: Applied Machine Learning, Causal Inference, Probability Theory, Quantitative Bias Analysis	
<i>Master of Public Health (MPH), Global Epidemiology (GPA 4.00/4.00)</i>	May 2020

• Inductee, Phi Chapter, Delta Omega Honorary Society (top 10% of students)	
• Global Field Experience Financial Award (\$2,250)	
• Charles C. Shepard Scholarship (50% tuition and \$2,000 summer practicum award)	

<b>Yale University</b>	<b>New Haven, CT</b>
<i>Bachelor of Arts (BA), History of Science, Medicine, and Public Health (GPA 3.48/4.00)</i>	May 2014

## ANALYTICS EXPERIENCE

<b>Epidemiologic Research &amp; Methods</b>	<b>Atlanta, GA</b>
<i>Epidemiology Consultant</i>	October 2021 – Present
• Conceptualized analytical studies involving exploratory data analysis and regression modeling of publicly available data of infectious diseases for biotechnology firms, ensuring full and clear documentation of all processes	

• Conducted literature reviews on published scientific literature and write reports for pharmaceutical clients	
<b>Emory University</b>	<b>Atlanta, GA</b>
<i>Graduate Student Researcher</i>	
October 2018 – Present	

• Collaborated with multiple study teams on infectious disease research projects, resulting in: 14 conference abstracts, 3 first-author papers, 3 co-authored papers, and 3 National Institutes of Health (NIH)-funded grants	
• Applied skills: R, SAS, Python, git, Zotero, EndNote, NVivo, Microsoft Office Suite (Excel and Powerpoint), statistical analysis, regression, survival analysis, longitudinal analysis, infectious disease modeling	

## BUSINESS EXPERIENCE

<b>Atlas Research</b>	<b>Washington, DC</b>
<i>Senior Analyst</i>	January 2018 – June 2018
• Engaged key stakeholders (Veterans, caregivers, and leaders of the Veterans Service Organization) in strategy meetings for the U.S. Department of Veterans Affairs (VA) to support Veterans who want to receive care at home	

• Managed project deadlines and drafted key deliverables, including one-pagers, slide decks, agendas, runs of show, and discussion questions for VA-hosted external meetings with key stakeholders	
----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	--

<b>Whiteboard Advisors</b>	<b>Washington, DC</b>
----------------------------	-----------------------

*Research Associate*

July 2016 – December 2017

- Leveraged communications, policy, research, management, and influencer engagement expertise to advise startups, corporations, and associations on strategy, ultimately driving sales
- Performed due diligence research projects for private equity firm clients and corporate clients interested in education technology, health and wellness, and corporate social responsibility (CSR) initiatives

**Global Health Strategies**

*Consultant*

**New Delhi, India**

July 2015 – October 2015

- Organized a fact-finding trip to Jakarta, Indonesia on immunization involving stakeholders in federal and local government, medicine, and advocacy for a Bill & Melinda Gates Foundation-funded project
- Led the development of a successful Gavi Request for Information for technical assistance in India and Kenya

# Santosh Tiwari

tiwarisantosht24@gmail.com | (650) 954-0236 | [Portfolio Website](#) | [LinkedIn Profile](#)

## WORK EXPERIENCE

### Genentech

Data Analyst Intern

South San Francisco, CA (Hybrid)

Jun 2023 - Sep 2023

- Built and revamped 5 interactive Tableau dashboards (Tours, Amenity, LobbyOps, CampusServices, Transportation) within 2 months enhancing real-time insights.
- Optimized data retrieval and visualization processes by leveraging SQL expertise and utilizing Amazon Redshift/dBeaver tools, ensuring seamless dashboard performance.
- Collaborated with cross-functional teams to develop and implement a streamlined process for collecting and analyzing customer feedback data, resulting in a 15% reduction in response time for addressing customer concerns.
- Developed and executed a comprehensive data tracking system, including 22 KPIs, to monitor key performance metrics, resulting in a 25% increase in overall project success rate.

### ICDR International

McLean, VA (Remote)

Digital Insights Specialist

Jun 2020 - Feb 2021

- Collaborated with web developers to implement design changes and optimize website performance, resulting in a decrease in bounce rate by 10%.
- Utilized web analytics tools such as Google Analytics to track website traffic and social media metrics, and analyzed data to identify patterns, resulting in a 30 % increase in website traffic and social media engagement.
- Analyzed audience engagement metrics and identified trends that led to the creation of targeted marketing campaigns, resulting in a 15% increase in customer conversion rate.

### One Degree Inc

San Francisco, CA (Remote)

Volunteer Data Entry Assistant

Apr 2020 - Jun 2020

- Executed comprehensive data cleaning and integrity checks, identifying and resolving errors in the company database, and ensuring top-level data quality.
- Verified and updated 200+ existing resources to ensure accuracy, utilizing error analysis and conducting additional research, resulting in a maintained accuracy rate of 98%.
- Implemented new data entry procedures in collaboration with the data management team, resulting in a 20% increase in efficiency and a 10% reduction in errors.

## EDUCATION

### San Jose State University

San Jose, CA

Master's, Data Science

Graduation Date: May 2025

### St. Cloud State University

Saint Cloud, MN

Bachelor's, Data Science (GPA: 3.47)

Graduation Date: Dec 2022

## PROJECTS

### Sales Forecasting for Walmart Stores

Oct 2022 - Nov 2022

- Utilized statistical techniques such as regression analysis and time series forecasting to identify key factors influencing sales trends with an accuracy rate of 95%.
- Generated accurate revenue forecasts for the company by analyzing historical sales data, market trends, and customer behavior, leading to an improvement in forecast accuracy.

### Impact of COVID-19 on Global Trade (Award Winning Project)

Mar 2022 - April 2022

- Conducted comprehensive data analysis using Tableau and R to identify and compare pre-pandemic vs pandemic trade value trends for eight countries across 3 continents, achieving a 95% confidence level accuracy.
- Developed and implemented an ARIMA model in Python to predict 2022 imports and exports, resulting in a forecast accuracy of over 90%.

### Instacart Market Basket Analysis

Dec 2021 - Jan 2022

- Developed and implemented machine learning algorithms using Python's sci-kit-learn library to identify the most frequently purchased items, resulting to an increase in sales for high-demand products.
- Analyzed customer purchase data and identified key product categories with the highest growth potential, leading to a recommendation for Instacart to expand its selection in those areas for an increase in revenue.

### Database System for a Sailboat Company

Aug 2021 - Dec 2021

- Designed and executed a comprehensive database system utilizing Oracle SQL, resulting in improved data management for sailboats, customers, orders, and inventory.
- Developed and executed advanced SQL queries to extract and analyze data from the Oracle database, resulting in a increase in data accuracy and efficiency.

## SKILLS

Python, R, SQL, PostgreSQL, SAS, C++, Tableau, Microsoft PowerBI, Microsoft Excel, ggplot2, Adobe Analytics, Google Analytics, Matplotlib, JMP, Amazon Redshift, DBeaver, AWS, Hadoop, Jira, Asana, Trello

## RELEVANT COURSEWORK

Statistical Modeling, Statistical Methods, Applied Probability, Applied Regression, Applied Categorical Analysis, Data Mining, Data Visualization, Data Analysis, Statistical Consulting, SAS Programming, Python Programming, R Programming, Time Series, Database Design.

## AWARDS AND ACTIVITIES

Math/Statistics/Computer Science Scholarships, Saint Cloud State University  
2022 Minnesota Data Winner, Advanced Level, Minnesota IT Centre of Excellence

Oct 2022

Apr 2022

Statistics and Math Tutor, College of San Mateo  
International Student Ambassador, College of San Mateo

Aug 2019 - Dec 2020

Aug 2019 - Dec 2020

# HANSOO HUDSON CHANG

PhD Researcher in Applied Cognitive and Brain Sciences | [aychsi.github.io](https://aychsi.github.io)

(814) 232 4335 | [hnssoo.chang@gmail.com](mailto:hnssoo.chang@gmail.com) | [github.com/Aychsi](https://github.com/Aychsi) | [linkedin.com/in/hansoo-chang](https://linkedin.com/in/hansoo-chang)

PhD researcher using statistical and machine learning techniques to predict mental health and mortality outcomes. Experience with solving pharmaceutical and technology industry issues backed by presentable scientific methods. Special expertise in developing neural nets, time-series analysis, Bayesian network analysis, multilevel linear regression frameworks, and natural language processing to develop advanced quantitative solutions in medical and consumer technology.

## RELEVANT EXPERIENCE

2021 - Present

PhD Researcher | Quantitative Statistics and Psychology Lab | Drexel University

- Designed advanced statistical frameworks (mixed-effects modelling, ggplot, XGBoost, Apache SQL) to forecast mental health and mortality risks over time in large electronic health records (Pediatric HIV/AIDS Cohort Study). In press for *NeuroImage*.
- Implemented data reduction techniques in big brain imaging data (Adolescent Brain and Child Development study) using Bayesian Network Analysis and time series modelling to classify biomarker trends. In press for *Statistics in Medicine*.
- Implemented behavioral predictive analytics (Decision Trees, NLP) in smartwatch app software to predict binge-eating episodes with Promptworks LLC achieving a binge-episode detection rate of above 80%.

2019 - 2020

Data Science Intern | Delve Inc.

- Delve Inc. is a product design consulting firm specializing in medical and consumer products.
- Summarized and suggested improvements to Samsung America regarding Samsung Flip and Fold phones based on large customer survey questions using Market Mix Modeling and sentiment analysis techniques in Python (AWS).
- Designed usability experiments and presented findings for Dexcom G6 CGM to Dexcom engineers using usability scores and A/B testing to significant improvements in new CGM design.

2020

Research Intern | Center for Healthcare Quality and Analytics (CHQA) | Children's Hospital of Philadelphia

- Re-modelled Scope Processor System used by nurses at CHOP to eliminate misplaced scopes based on existing text-based electronic health records using NLP (NLTK) in Python.
- Eliminated misplaced scopes by 75% in first 6 months of re-design.

## PERSONAL PROJECTS

MLB Wins-Loss Predictor Model

- Designed a novel wins-loss model for Major League Baseball teams using K-Means clustering and Regression techniques. Model differed from actual results by 2.687 wins on average, beating most conventional wins-loss predictor algorithms on ESPN.

Board Game Betting Model and Simulator

- Used Python to implement a virtual version of popular board game "Camel Up," which is a horse-betting board game.
- Implemented algorithm to calculate most optimal betting strategy based on game-state. Utilized to beat competitors at local board game café.

## EDUCATION

2021 – Present

PhD in Applied Cognitive and Brain Sciences | Drexel University | Philadelphia, PA

Advisor: Fengqing Zoe Zhang, PhD | GPA 4.0 / 4.0

7 peer-reviewed journal papers | 2 first author journal papers | 4 Conference Presentations

2017 – 2021

BA in Cognitive Science, Minor in Computer Science | University of Pennsylvania | Philadelphia, PA

Advisor: Charles Yang, PhD | GPA 3.5 / 4.0

## SKILLS

R | Python | Machine Learning | Tableau | Microsoft Azure | XGBoost | Spark | ggplot | NLP | Excel | AWS | SQL |  
Industry Research | Naïve Bayes| Decision Trees | Predictive Analytics

11/03/2023

To whom it may concern,

I am writing to express my genuine interest in the Applied Machine Learning Intern position. With a diverse background in statistics, machine learning, product design, and business analytics, I believe that I am the ideal candidate for this position and am eager to bring my expertise to this role.

Currently, as a PhD Researcher at Drexel University in the Applied Cognitive and Brain Sciences program, my endeavors in the realm of machine learning have culminated in impactful outcomes. Specifically, my academic journey afforded me the opportunity to independently design advanced statistical frameworks, utilizing tools like mixed-effects modelling, XGBoost, and Apache SQL. These frameworks were instrumental in my research for forecasting mental health and mortality risks within the Pediatric HIV/AIDS Cohort Study, research which has been recognized and is currently in press for *NeuroImage*. My exploration into brain imaging data, specifically within the Adolescent Brain and Child Development study, saw me employ data reduction techniques such as Bayesian Network Analysis to examine the relationship between abnormal brain development and certain medications. This endeavor aimed at classifying biomarker trends, further solidifying my reputation in the field, with my findings soon to be published in *Statistics in Medicine*.

My accomplishments within academia have been applicable in industry as well. My collaborative efforts with various companies, including my consultancy role for Modus Create LLC, and my personal projects are reflective of my entrepreneurial spirit. As a consultant, I implemented behavioral predictive analytics using Decision Trees and Support Vector Machines in a smartwatch app software, achieving the innovative goal of predicting binge-eating with over 80% accuracy. I pride myself on my adaptability and my ability to seamlessly integrate with diverse teams, always aiming to bridge the gap between data science and actionable business insights.

Furthermore, as a Data Science Intern at Delve Inc. and the Children's Hospital of Philadelphia, I successfully analyzed and proposed improvements to a variety of life-changing products. For Samsung America's Galaxy Fold and Flip phones, I leveraged advanced text-scraping and sentiment analysis techniques to provide insights into customer satisfaction on their new phones and its effect on sales. Moreover, I designed and presented usability experiments for Dexcom's flagship G6 CGM product, employing rigorous data analysis methodologies such as Marketing Mixed-Effects Linear Regression and A/B Testing to verify that this new model improved over the last model. These experiences mirror the primary responsibilities outlined for your role, wherein deriving product solutions through data analysis, and communicating the findings to business leaders is paramount.

In conclusion, I am confident that my strong academic background from the University of Pennsylvania and Drexel University, combined with my interdisciplinary work experiences, align perfectly with the requirements and ethos of your position. Thank you for considering my application. I look forward to the possibility of contributing to Merck and am available at your earliest convenience for a discussion.

Sincerely,  
Hansoo Chang

# Jamal Ansary

Toledo, Ohio

(419) 340 - 9877 ◊ Jimi.ansary@gmail.com

[Google Scholar](#) | [Github](#)

## Experience

### University of Toledo

*PhD Research Assistant*

**Aug 2021 – Present**

*Toledo, OH*

- Working on NSF funded project to study the applications of **Deep Learning** and **data analysis** in membrane design
- Building a LSTM-based **Neural Network algorithm** for forecasting membrane performance, achieving acc of **91%**
- Addressing the problem of performative prediction in **time-series** data through optimization and **statistical analysis**
- Developing **Anomaly Detection** algorithm for multivariate time series data, increasing the acc by **15%**

### Speed of Me

*Machine Learning Engineer (Internship)*

**May 2023 – Aug 2023**

*San Francisco, CA*

- Achieved **83%** accuracy in predicting customer bandwidth by developing a **Deep Learning** pipeline using **TensorFlow** and historical network data on a large deadset (300k)
- Developed an internet provider **recommendation system** based on costumer location and band-with requirement
- Performed data processing, data cleaning and visualisation to represent visitors segments, enabling targeted marketing

### University of Toledo

*M.Sc. Research Assistant*

**Jan 2018 – Aug 2021**

*Toledo, OH*

- Developed a hybrid **Deep Convolution Neural Network** for real-time detection of algae blooms in water using **aerial images** with accuracy of **85%** with **TensorFlow and OpenCV** (2k images)
- Built an augmentation pipeline, resulting in **15% reduction** in false positive detection rate in **CNN, SSD**
- Designed an interactive web platform for deploying deep-learning models using **AWS and Google Cloud**
- Collaborated on a machine learning project for predicting COVID-19 patients discharge rate through applying survival analysis methods, achieving acc of **71%** (**covered in local news**)

## Technical Skills

**Languages:** Python, MATLAB, SQL, L<sup>A</sup>T<sub>E</sub>X

**Machine Learning Libraries:** Tensorflow, PyTorch, Keras, Scikit-learn, Numpy, Pandas

**Tools and Platforms:** Google Cloud, AWS, Ohio Supercomputer, VS Code, Git

**Visualization Tools:** Tableau, Bokeh, Matplotlib, Plotly

**Database:** MySQL, SQLite

## Education

### University of Toledo

*Ph.D. in Mechanical Engineering*

**Aug. 2021 – Dec 2024 (Expected)**

*Toledo, OH*

### University of Toledo

*M.Sc. in Mechanical Engineering*

**Jan. 2018 – Aug 2021**

*Toledo, OH*

### Azad University

*B.Sc. in Electrical Engineering*

**Aug. 2011 – Dec 2016**

*Tehran, Tehran*

## Publication

- **Ansary, Jamal**, et.al. "Comparative Performance of Traditional and Neural Network Time Series Forecasting Models in Predicting Membrane Sustainability and Distillation Performance" (under review). ACS Journal (IF=16.3)
- Nemati, Mohammadreza, **Ansary, Jamal**. "Machine-learning approaches in covid-19 survival analysis and discharge-time likelihood prediction using clinical data". Cell Patterns (IF=6.5)
- **Ansary, Jamal**, et.al. "Swarm of USV, from simulation to real world implementation. International Design Engineering Technical Conferences"(IDETC-CIE)

## Honors and Awards

Full academic scholarship for PhD and master studies at University Of Toledo

COVID-19 research for developing Survival Analysis model for COVID-19 patients was covered at **local news channels**

Cloud Computing resources award recipient from **Ohio Supercomputer** and Path from **NSF**

Best TA award recipient at College of Engineering at University of Toledo

First Place at research presentation competition in Mechanical Engineering Department at University of Toledo

Advisor for senior Capstone projects at University of Toledo

Vice President of Persian Student Organization at University of Toledo

[sonyajin@stanford.edu](mailto:sonyajin@stanford.edu)  
Palo Alto, CA

# Sonya Jin

(609) 937-6888  
<https://www.linkedin.com/in/sonya-jin/>

Stanford CS/AI student with skills & experiences in leading/engineering the development of ML-centric MVPs, with a focus in healthcare.

## EDUCATION

**Stanford University** Graduating Year: 2025

**Computer Science with Minor in Symbolic Systems - Biomedical Applications, Artificial Intelligence**

GPA: 3.8

Relevant Courses: Programming Methodology/Abstractions, Building for Digital Health, Principles of Data Science, Machine Learning in Neuroimaging (Grad/PhD), Probability for Engineers, Biodesign for Digital Health

**Organizations:**

ASES Global Entrepreneurship - Breakthrough Venture Capital Subteam

Stanford Students in Biodesign (SSB) - Stanford Health Innovations in Future Technologies (SHIFT) Subteam, **Project Lead**

## EXPERIENCE

**Amazon Web Services (AWS)** July, 2023 - September, 2023

*Applied Scientist, Intern - Graph Machine Learning*

- Spearheaded contrastive learning research with graph group discrimination (GGD) that supported the product development and implementation feasibility of heterogeneous graph functions in Deep Graph Library (DGL), a Python package with 769 citations and 10,000+ users geared to make graph ML research more efficient and scalable.
- Employed multi-relational learning techniques and streamlined data preprocessing using PyTorch and Deep Graph Library (DGL) for heterogeneous graphs; approach improved performance by 26% when compared to graph contrastive learning benchmarks.

**Stanford Byers Center in Biodesign** January, 2023 - March, 2023

*Technical Product Manager / Lead Engineer*

- Co-led agile development of telehealth patch testing app, PatchTrackr, that reduced annual care costs by 50% and decreased waiting times by 60% for allergic contact dermatitis patients. Prioritized requirements for app to be deployed at Stanford Hospital Dermatology Department in two months.
- Led feature development, including augmented reality (AR) patch detection and EHR integration, with a cross-functional team of 8; engineered ARCamera feature; **won First Place at the Winter '23 Building Digital Health Cohort Pitch Competition.**

## PROJECTS

- **Dynamic Health - Precision At-Home Physical Therapy**
  - Leading development of a pose detection telehealth app for personalized, guided, precise at-home physical therapy with patient-defined benchmarks and visualized progress to increase patient adherence and recovery efficacy. Currently in MVP development - scheduled for pilot studies at on-campus PT clinics with student-athletes next year.
- **NurtureU - Platform for Streamlining Transition of Care with International Data Interoperability**
  - Collaborating with a cross-functional team to smoothen transition of care for international students who are unfamiliar with US healthcare; empowers students to have autonomy over portable, digitized data with translation support.
- **NanoGen: Vaccine AI Research Tool**
  - Engineered an app to predict sequence degradation given an mRNA sequence, expediting discovery of stable mRNA vaccines. **\*Won 3rd place at the International NucleoTIDE Biotechnology Hackathon**
- **Expressiva: AI EDM Production Platform**
  - Streamlined EDM production workflow with automated raw audio to MIDI conversion and melody/accompaniment completion with deep learning. **\*Won Best Far Out Hack at Creatica Hackathon**

## SELECT AWARDS

- Regeneron Science Talent Search (STS) Scholar 2022, Named America's Top 300 Young Scientists
- Semifinalist, Team Captain - Modeling the Future Mathematical Modeling Competition
- Intel International Science & Engineering Fair - 1st Place in Computer Science in Delaware/Pennsylvania/New York Regions

## SKILLS

Product Management: Agile MVP development, KPI definition/data analysis, requirement prioritization

Technical: Python, C++, HTML, CSS, Tensorflow, DGL, PyTorch, Swift

## Veronica Jenkins

State College, PA • 267-272-2148 • [vfj5053@psu.edu](mailto:vfj5053@psu.edu) • [www.linkedin.com/in/veronicajenkins1](http://www.linkedin.com/in/veronicajenkins1)

### Education

The Pennsylvania State University B.S. Computational Data Sciences	Expected May 2025 GPA: 3.59
-----------------------------------------------------------------------	--------------------------------

### Skills

•Python (object oriented programming)	•SQL	•Data visualization
•R (supervised machine learning, data analysis)	•Research	•Critical thinking

### Research Experience

#### Department of Homeland Security Center of Excellence: Center for Accelerating Operational Efficiency (June 2022)

##### *Research Intern*

- Developed an optimization problem using operations research to determine the best strategy to staff a TSA checkpoint
- Created presentation outlining improvements of accessibility of airports to solve various issues
- Utilized NetLogo, Excel and Tableau to create various models and visualizations
- Won first place out of 6 groups in SEQAL Design Competition

#### NASA Pennsylvania Space Grant Consortium Research Internship Program (January 2022-May 2022)

##### *Research Intern*

- Collaborated with graduate student in the Cockburn Laboratory at Pennsylvania State University mentor to design primers and plan research project
- Executed polymerase chain reactions, restriction digestions, and ligation reactions

### Extracurricular Experiences

#### Association of Women in Computing (2021-Present) Director of Student Relations

- Help foster a welcoming community for women and non-binary computing students at Penn State University
- Execute ideas to help women find success in computing
- Create events for over 50 students to attend to help increase retention of women in computer science

#### Girls Who Code Summer Immersion Program (July 2021) Participant

- Utilized HTML, CSS, and JavaScript to create a website to encourage environmental justice activism

#### Girl STEM (2017-2021) Youth Instructor + Member

- Taught science lessons and hosted science clubs at the local elementary school to encourage youth interest in STEM

#### Interact Club (2017-2021) Coordinator

- Helped organize and participated in fundraisers, baked dog treats to sell for Roxy reading and therapy dogs, participated in Feed Our Starving Children food packing event

#### Rotary Youth Exchange Scholarship (2019-2020). Exchange Student

- Selected to receive scholarship to live in Brazil
- Learned Brazilian Portuguese to effectively communicate with locals and embrace culture

#### Pearl S. Buck Global Leadership Program (2018) Financial Committee Head + Participant

- Led Financial Committee of an event that raised \$1,300 for orphanages in Vietnam to improve living conditions, worked collaboratively to maintain budgets

# Mason Kim

[mmh.kim@gmail.com](mailto:mmh.kim@gmail.com) | 972-730-1050 | <https://www.linkedin.com/in/masonkim1/>

## EDUCATION

---

### Carnegie Mellon University, H. John Heinz III College

*M.S. in Healthcare Analytics and Information Technology*

*May 2025*

- **Coursework:** Econometrics, Finance, Management Consulting, Health Systems

*B.S. in Statistics and Machine Learning, Minor in Business Analytics and Optimization*

*May 2023*

- **Coursework:** Probability, Data Structures/Algorithms, ML, Applied ML, Regression Analysis, Database Management

## SKILLS

---

**Programming Languages:** Python, R, SQL, HTML, CSS, JavaScript, Java, C

**Visualization Tools:** Tableau Desktop, Tableau Prep Builder, Excel

## PROFESSIONAL EXPERIENCE

---

### Optum - UnitedHealth Group, Pittsburgh, PA

*Technology Development Intern – Data Science Concentration*

*Jun 2022 – Aug 2022*

- Performed exploratory data analysis on patient data through Jupyter notebook to investigate general patterns across the population
- Predicted probability of a patient contracting sickle cell and thalassemia utilizing patient datasets with over 1 million entries and the popularized RETAIN model (a modified LSTM)
- Responsible for data manipulation, data cleaning, and tuning the hyperparameters of the model using Pandas and PyTorch framework to improve the model's time complexity and accuracy
- Led and participated in multiple data science, software, and product management meetings and presented bi-weekly at a national conference to executives and product managers call to discuss model's progress

### Pittsburgh Supercomputing Center, Pittsburgh, PA

*Undergraduate Research Programmer*

*Jan 2022 – Apr 2022*

- Assisted in designing a machine learning algorithm that predicted the queue time for XSEDE Supercomputer and conducted research by analyzing similar, existing models

### involveMint, Pittsburgh, PA

*Data Analyst Intern*

*Jul 2021 – Aug 2021*

- Designed a dashboard in Python that displayed both basic and advanced models, accepting a variety of parameters for the business development team to adjust and utilize at their stakeholder meetings
- Reduced data collection process from 30 minutes to real time by updating the previous SQL scripts and connected the database directly to the dashboard

## RESEARCH EXPERIENCE

---

### Carnegie Mellon University, Pittsburgh, PA

*Graduate Research Assistant, iLab*

*May 2023 – Present*

- Developed and refined a R Shiny, data science application for physicians and patients that classifies the risk level of heart failure and signifies important variables given clinical datasets of size 10K+
- Improved the performance of the PCA-LDA model by increasing the AUC score from 0.7 to 0.85 to help the model distinguish between classes more effectively
- Integrated multiple comparison features and miscellaneous data visualizations for users including clustering and regression models through R to provide more valuable insight regarding the patient's probability of contracting cardiovascular disease

### Research Programmer, Dynamic Decision-Making Laboratory

*Aug 2022 – Dec 2022*

- Developed a 10,000-line, full-stack Django web application to build a web game that educated users on the security risks of cyber attacks
- Led weekly software meetings, presented code to lab members, and discussed research implementations of the app

## ACADEMIC PROJECTS

---

### Aphasia Project

*Research Capstone*

- Improved the reliability and validity of the comprehensive measurements of fluency that helps detect patterns of speech in Aphasia patients by performing statistical analysis on clinical trials
- Demonstrated statistical and interpersonal skills and directly applied concepts to a real-world healthcare problem through consistent communication with the advisor and weekly presentations
- **Conference Presentation:** Fromm D, MacWhinney B, Chern S, Geng Z, Kim M, Greenhouse J. Automated Analysis of Fluency Behaviors in Aphasia. Poster Presented at: Clinical Aphasiology Conference 2023; 2023 May 30-Jun 03; Atlantic City, NJ. [[Link](#)]

# **Yunqing Qiu**

400 W 113th St Apt 427, New York, NY

614-906-6440 | yq2374@columbia.edu | <https://www.linkedin.com/in/yunqing-qiu/>

## **EDUCATION**

### **Columbia University**

*M.S. in Data Science, GPA: Not Yet Available*

New York, NY

expected Dec 2024

Courses: Probability and Statistics, Exploratory Data Analysis and Visualization (R), Natural Language Processing.

### **The Ohio State University**

*B.S. in Computer Science and Engineering, GPA: 3.9/4.0*

Columbus, OH

May 2023

Courses: Machine Learning, Database (SQL), Data Structures and Algorithms, Software Development and Design (Java), Operating Systems, Computer Networking, Principles of Programming Languages (C).

## **PROFESSIONAL EXPERIENCE**

### **Tencent**

Shenzhen, China

*Technical Product Manager Intern*

May 2021 - Aug 2021

- Conducted competitive strategies analysis of an AI Content Creation product to differentiate product cost-benefits, documented and presented a 30-minute PPT on product positioning and market analysis at department meetings.
- Designed flow diagrams and UI prototypes for image function, planned and coordinated with engineer teams, and finished development in 6 weeks; analyzed performance with A/B Test feedback to help enhance later iterations.
- Employed a web data collector to obtain 10k+ content labels to facilitate the Content Learning and Multimodal Understanding work; automated label-collecting task and promoted efficiency for algorithm team.

## **PROJECTS AND RESEARCH**

### **Research on Media Effects Psychological state via PokéMon GO**

Sep 2022 - Dec 2022

- Conducted literature review; summarized and examined mobile data pre-processing methods on 1000+ samples.
- Extracted and integrated unstructured sensor data with SQLite database, cleaned and performed EDA on self-report data using Pandas in Python, and executed feature selection on self-report indicators for modeling.
- Applied Agile Methodology to project management, and saved 4 days for the 2<sup>nd</sup> iteration versus primitive plan.

### **Business Analysis of Healthcare Data Project**

Sep 2022 - Dec 2022

- Executed preprocessing on the Kaggle heart disease dataset; utilized KNN, random forest, and SVM classifiers with Scikit-Learn and tuned hyperparameters to evaluate optimal classifiers and predict maximal benefit.
- Generated easy-to-read conclusions with visualized graphs to show accuracy, tendency, and limitation; delivered business strategy recommendations in a concise 3-page pre-read report and 5-minute presentation.

### **Research on Rice Yield Prediction Based on LSTM and GRU**

Oct 2022 - Nov 2022

- Performed EDA and pre-processing including missing values handling, one-hot encoding, and time-series dataset reformat on 130,000 pieces of high-dimensional meteorological data.
- Designed experiments for a PyTorch framework to test the impacts of depths and direction on several NLP models.
- Developed and optimized a GRU model to predict grain yield, reducing test loss by 92-98%. Authored and published paper Rice Yield Prediction Based on LSTM and GRU, is indexed by EI Compendex.

### **Student Graders Web Application Development Project**

Jan 2022 - May 2022

- Programmed, tested, and deployed a web student graders management system based on the Ruby on Rails framework, realizing functions of query courses, request and assign graders, 3 role differentiation, etc.
- Served as shift project manager, leading the first 50% of development, including demand analysis, database design, page design, construction, data crawling, implementation, and debugging of partial functions.
- Acquired top 1 overall scores based on functional design, user experience, and appearance in class.

### **Mental Health Hack Fest: A Data Visualization Competition**

Apr 2021

- Pre-processed and explored data from 2,000+ participants on the COVID-19 mental health survey with 50 questions with Excel; used R to visualize graphs and discovered primary factors and proportions affecting mental health.
- Won the Community Thinker Award; assigned the \$500 (naming right) reward to the Asian Mental Health Collective.

## **SKILLS, TECHNOLOGIES, AND LANGUAGES**

Languages: Python (NumPy, Pandas, SciPy, PyTorch, Matplotlib), SQL, R (GGPlot), Java, C, JavaScript, HTML/CSS.

Software and Dev. Tools: Microsoft Excel/PowerPoint/Word, Tableau, Jupyter Notebook, Google Colab, Linux, Git.

# Nithin Parthasarathy

---

16 Mahogany Drive, Irvine, CA 92620 • [nithinparthasarthy@gmail.com](mailto:nithinparthasarthy@gmail.com) • 949-878-1989

## Education

<b>UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN</b>	Urbana, IL
B.S., Major: Computer Science, Minor: Mathematics, GPA: 4.0	May 2025
<ul style="list-style-type: none"><li>• Awards: Chancellor Scholar, James Scholar, Engineering Visionary Scholar, Dean's List</li><li>• Relevant Coursework: Data Structures (Honors), Algorithms, Artificial Intelligence, Computer Architecture, Probability and Statistics, Linear Algebra</li></ul>	

## Experience

<b>Apple Inc.</b>	Champaign, IL
Campus Leader	May 2023 - Present
<ul style="list-style-type: none"><li>• Organized Apple device workshops and managed relations with Apple partner campus organizations. Collaborated with other Campus Leaders to develop Swift apps.</li></ul>	
<b>FOCAL Lab at UIUC</b>	
Researcher	Champaign, IL
<ul style="list-style-type: none"><li>• Developed effective adversarial attack algorithms that are introduced into novel Brain Computer Interface systems, called EEGNet, to assess their robustness.</li></ul>	May 2023 - Present
<b>UCLA, Department of Radiological Sciences</b>	
Researcher	Los Angeles, CA
<ul style="list-style-type: none"><li>• Developed new brain-computer interfaces (BCI) using language models and created word completion algorithms, providing a 40% improvement over currently deployed systems. Published work in peer-reviewed IEEE/BCI society conferences.</li></ul>	March 2019 – Present
<b>Emotiv</b>	
Summer Intern	San Francisco, CA
<ul style="list-style-type: none"><li>• Analyzed Neurotechnology-based wellness experiments with Dell to provide feedback to employees on how to personally optimize their breaks, improving their focus/efficiency.</li></ul>	June - August 2022
<b>Neurable</b>	
Summer Intern	Boston, MA
<ul style="list-style-type: none"><li>• Improved ML algorithms by suppressing noisy electrodes in Enten headphones to increase user productivity based on the user's brain signals.</li></ul>	June - August 2021

## Leadership & Activities

<b>Zero-Waste Initiative</b>	Irvine, CA
Founder and President	2020 - Present
<ul style="list-style-type: none"><li>• Started a volunteer-based non-profit, providing baked goods to different charities serving the local community. Over \$250,000 worth of baked goods supplied to charities. Recognized by the US Congress. Won 7 international awards.</li></ul>	

## Skills & Interests

**Technical:** C++, Java, Python, MATLAB, Verilog, MIPS, Swift, NLP, TensorFlow, PyTorch  
**Interests:** Community Service, Indian Flute, Western Flute, Piano, Tea, Fantasy Novels

# Shubham Sudhir Gurav

SYRACUSE NY 13210 | +1 315 952 7339 | ssgurav@syr.edu | [www.linkedin.com/in/shubham-gurav/](https://www.linkedin.com/in/shubham-gurav/)

## EDUCATION

Syracuse University   M.S in Computer Science (Machine Learning Concentration)   GPA 3.5/4.00   Syracuse, USA	May 2024
Mumbai University   Bachelor of Computer Engineering   GPA 8.52/10.00   Mumbai, India	June 2020

## PROFESSIONAL EXPERIENCE

<b>Machine Learning Engineer Intern   ESG Flo   Brooklyn, USA</b>	Sept 2023 – Present
<ul style="list-style-type: none"><li>Developed an <b>online streaming clustering algorithm</b> utilizing SOTA CLIP model for embedding generation and document clustering, enabling the rapid categorization of similar documents, resulting in a reduction in manual effort and time.</li><li>Established a robust machine learning architecture pipeline to streamline 1TB+ ESG data for the sustainability improving the efficiency of document processing by 40% while maintaining a clustering accuracy rate of over 90%.</li></ul>	
<b>Machine Learning Research Intern   Comcast Applied AI   Philadelphia, USA</b>	May 2023 – Aug 2023
<ul style="list-style-type: none"><li>Enhanced Xfinity Chatbot's Annotation accuracy by using state-of-the-art supervised <b>representation learning</b>, to understand the contextual meaning and similarity of sentences, resulting in 80% improvement in classification accuracy for user intents.</li><li>Achieved up to 40% improvement in downstream tasks, including <b>Clustering</b> and <b>Information Retrieval</b>, by developing models that generate meaningful custom sentence embeddings using Sentence Transformer &amp; MP-Net.</li></ul>	
<b>Machine Learning Engineer   Tata Consultancy Services   Mumbai, India</b>	Aug 2020 - Aug 2022
<ul style="list-style-type: none"><li>Worked on NLP for Information retrieval, optimizing question-answering components from legal documents.</li><li>Fine-tuned Roberta for question answering to extract key datapoints from legal documents, reduced false positives by 3x over conventional spaCy pattern matching, leading to production ready code.</li><li>Trained Text-to-Text Transformer to correct OCR post-processing grammar, resulting in 40% improved extraction scores.</li><li>Deployed T5 for inference on CPUs using quantization and ONNX runtime, improving inference speed by 3.5x.</li><li>Successfully onboarded 2 trainees through KT sessions and presentations to align them with company culture and vision.</li></ul>	
<b>Artificial Intelligence Intern   IITM   Pune, India</b>	June 2019 - June 2020
<ul style="list-style-type: none"><li>Collaborated with NASA and researched state-of-the-art deep learning algorithms such as CNNs, 2D Image reconstruction to generate high-resolution (4x) downsampled rainfall data, reducing Root Mean Square Error by 25%.</li><li>Perform large-scale distributed training of deep neural networks using HPC and dedicated clusters, resulting in upscaling of spatial accuracy and temporal precision by 45%.</li></ul>	

## SKILLS SUMMARY

- Programming Tools:** Python, C, Haskell, Pytorch, TensorFlow, Keras, Pandas, NumPy, Scikit-Learn, Flask, Spark, SQL.
- Domain Expertise:** AI, Machine Learning, Deep Learning, Computer Vision, NLP/NLU, ML Ops, Kubernetes, Image Processing, CNN, RNN, Unit/Regression Testing, Distributed Computing, Transformer-based language modeling.
- Cloud technologies:** AWS, Azure ML Studio, AWS Sage Maker, MLflow, GIT, Databricks, Docker, Gitlab, Jira, Hadoop, Spark.

## RESEARCH PUBLICATION

- Paper titled “Deep-Learning based down-scaling of summer monsoon rainfall data over Indian region” published in Theoretical and Applied Climatology Journal by Springer 2020.

## TECHNICAL PROJECTS

<b>Indian Space Research Organization (ISRO)   Mumbai, India</b>	Mar 2020 – Jun 2020
<i>Nowcasting for Severe Weather Prediction</i>	
<ul style="list-style-type: none"><li>Devised a Time series predictive model (LSTM) to accurately predict risks of severe weather with 82% accuracy under guidance from ISRO, using multichannel Geo-stationary INSAT-3D data from Satellite.</li><li>Integrated view of real-time and multi-dimensional images &amp; frames, incorporated prediction summaries into the model to alert and enable mitigation against severe weather with a benefit of 5X faster timeframes.</li></ul>	

## LEADERSHIP

<ul style="list-style-type: none"><li><b>Social Chair</b>   Electrical &amp; Computer Science Graduate Organization   Syracuse, USA</li><li><b>Founder</b>   Codex - Ed-Tech preparing students for professional journey   Mumbai, India</li><li><b>President</b>   Social Wing – A Non-Profit student organization   Mumbai, India</li></ul>	Aug 2022 – Present
	July 2021 – July 2022

## ACHIEVEMENTS

- Winner of National Level Smart India Hackathon competed against 100k participants from 25+ states in India.

# AKSHAJ KUMAR

---

akshajkumar.netlify.app

630-854-6141

akshajk8@gmail.com

Open to relocation

---

## SKILLS

Python, Java, R, MATLAB, C/C++, HTML/CSS, SQL, SAS, JavaScript, Swift, Git, Agile/Scrum, Power BI, Linux, Jira  
Libraries: Pandas, Numpy, Scikit-learn, TensorFlow, PyTorch, Matplotlib

## EXTRACURRICULAR ACTIVITIES

**Sun Devil Data Science/ Software Developers' Association (SoDA)/ Artificial Intelligence Club**

Attend technical workshops on critical data science tools such as Python, exploring the Software Engineering field, and developing the skills required to solve real-world problems with the application of AI

## ASU AI Scholar

Competed in a hackathon to design and develop a new skill for Amazon's Alexa

Led a five-member team of developers and designers to innovate upon Alexa's ability in education by using WolframAlpha's API

## EDUCATION

**Arizona State University**

**M.S. Computer Science**

GPA: 4.0/4.0

Relevant Coursework: Statistical Machine Learning, Data Processing at Scale, Data Visualization, Knowledge Representation and Reasoning, Software Verification/Validation/Testing

**B.S. Computer Science; Minor: Statistics**

**Tempe, AZ**  
**Jan 2023 - May 2024**

GPA: 3.6/4.0; Dean's List: Fall 2018, Spring 2020

Relevant Coursework: Data Structures/Algorithms, Database Management, Operating Systems, Applied Linear Algebra, Software QA and Testing, Business Intelligence

## EXPERIENCE

**Starbucks Corporation**

**Scottsdale, AZ**  
**May - Aug 2022**

**Software Development Engineering Intern**

- Extracted raw search data using REST API and Application Insights from Microsoft Azure running Elasticsearch
- Applied Natural Language Processing techniques to preprocess data, preparing it for training a machine learning model
- Implemented K-means clustering to predict related searches and achieved a significant 68% improvement in search accuracy for a user base of over 110,000 store and corporate employees

**American Express**

**Phoenix, AZ**  
**Jun - Aug 2021**

**Software Engineering Intern**

- Designed and built a headless API that integrated with Slack using Python/Flask to streamline post-COVID office transition for over 70,000 employees
- Utilized Microsoft Graph API to collect colleagues' availability data to display the best "in-office" availability for selected colleagues and enhance collaboration
- Ensured high availability and scalability by deploying the service to the cloud in a Docker container and scaling it with Kubernetes pods

**IBM Volunteer**

**San Francisco, CA**  
**May - Aug 2020**

**Business Transformation Consultant Intern**

- Gained hands-on experience with Javascript, Angular, HTML, and SCSS while experimenting with a supply chain demand prediction dashboard for Walmart
- Collaborated with an IBM partner to define and shape an initiative for Steward Health Care's go-to-market strategy, employing business process outsourcing (BPO) and robotic process automation (RPA) to reduce costs

**AJ Design Works LLC**

**Chicago, IL**  
**May - Aug 2019**

**Front-End Developer Intern**

- Developed a public-facing website for a medium-sized Realtor company using ReactJS and TypeScript, with WordPress as the underlying CMS
- Leveraged HTML/CSS to create two mass email newsletters, enabling a client to effectively reach out to a user base of 30,000 people through Mailchimp
- Performed extensive integration, regression, and compatibility testing to ensure seamless cross-device performance, using an agile workflow to iterate and improve the product

## PROJECTS

**Handwritten Digit Recognizer**

**May - Aug 2023**

- Implemented a Python deep learning model for handwritten digit recognition utilizing Keras and a convolutional neural network architecture, achieving an accuracy of 99.7% on the MNIST dataset

**Beat Maker**

**Aug - Nov 2022**

- Developed a Python-based beat maker using PyGame with a GarageBand-inspired UI, enabling users to create, save, and load custom drum beats

**Weather App**

**Jul - Sep 2021**

- Created a Python GUI weather application using Tkinter, Google V3 Geocoding API, and OpenWeatherMap API

**Snake Game**

**Jun - Aug 2020**

- Self-taught Python programming by developing the classic game of Snake as a desktop application using PyGame

**Robotic Car - Class Project**

**Oct - Dec 2018**

- Worked with three other students to design an autonomous robotic car that could navigate through a maze using MATLAB with second fastest maze navigation time among twelve teams

# Daniel Clepper

New York, NY | 630-200-4004 | dpc2147@columbia.edu | [LinkedIn](#) | [GitHub](#)

## EDUCATION

### Columbia University

New York, NY

#### Master of Science in Data Science

Anticipated Dec 2024

Selected coursework: applied machine learning, probability and statistics, algorithms for data science.

### University of Wisconsin-Madison

Madison, WI

#### Bachelor of Science in Industrial Engineering

May 2023

Graduated with Highest Distinction; Dean's Honor List (7/7 semesters)

GPA: 3.9/4.0

Selected coursework: operations research, decision science, machine learning, statistics, economics, linear algebra, data ethics.

## PROFESSIONAL EXPERIENCE

### Visa

Miami, FL

#### Data Scientist Intern

May 2023 - Aug 2023

- Developed custom genetic algorithm-based feature selection method resulting in 6.7% improvement in ROC-AUC for cross-border fraud detection XGBoost model trained on 100+ terabytes of transaction data.
- Wrote documentation on feature selection methods for 150-person predictive modeling organization.

### Branch Financial

Columbus, OH

#### Data Analyst Intern

Sept 2022 - Jan 2023

- Decreased incorrect marketing source attribution by 98% for users issued combination insurance policies through data tracing and SQL query development in BigQuery environment.
- Created and maintained 12 sales agency and growth strategy dashboards using SQL Server and Qlik Sense.

### Kraft Heinz

Chicago, IL

#### Business Analyst Intern

Jun 2022 - Aug 2022

- Pipelined data and prepared Tableau visualizations highlighting critical areas in enterprise architecture maturity and lifecycle to inform \$50MM+ annual technology spend across 1,000+ global business units.
- Built Power Automate workflow and framework to improve ease of use for enterprise architecture request process and IT leadership collaboration, reducing request processing time by 85%.

### Extreme Engineering Solutions

Madison, WI

#### Data Analyst Intern

May 2021 - Dec 2021

- Developed Python and SQL scripts to automate data extraction and improve efficiency of documenting changes to 650 controlled documents in company-wide quality management system, realizing savings of \$75K+ annually.
- Facilitated series of 5 test procedure BOM software rollouts utilizing agile project management methodology.

## PROJECTS

### Venture Capital Startup Success Prediction (Python)

- Built logistic regression and random forest models predicting success of private and public corporations using Crunchbase data, achieving an AUC score of 0.86 with a positive predictive value (PPV) of 97.9%.

### Gas Station Simulation Analysis (Arena)

- Performed sensitivity and alternative analyses on Costco gas station queueing data to identify an optimal queueing system with a 48% reduction in waiting time and 12% greater total output, on average.

### Global Life Expectancy Exploration (R)

- Conducted hypothesis tests and linear regression to report on and explain global differences in life expectancy.

## LEADERSHIP

### Wisconsin Engineering Student Council – Executive Board

### Emerging Leaders in Engineering – Project Manager, Leadership Advisory Council

### DotData Data Science Club – Datathon Event Team, General Member

## SKILLS

**Technical:** Python (Scikit-Learn, PyTorch, Pandas, NumPy, Matplotlib), Hadoop, Spark, R, SQL, Julia, Tableau, Excel

# Charlie Plonski

862-324-6880 • [charles.plonski@gmail.com](mailto:charles.plonski@gmail.com) • <https://www.linkedin.com/in/charlie-plonski-4505581b1>

## ELECTRICAL ENGINEERING • COMPUTER SCIENCE • MACHINE LEARNING• LEADERSHIP

Innovative Electrical Engineer who is Honest, Curious, and Committed to Continuous Improvement. I am interested in the combination of Signal Processing and Machine learning applied to IoT and sensor systems.

### Education

**B.S. Electrical Engineering, Minor in Computer Science – University of Illinois at Urbana-Champaign**

**GPA: 4.0/4.0, James Scholars Honors Program, Expected Graduation: May 2024**

**Received Provisional admission to ECE MS/PhD program at UIUC for Spring/Fall 2024**

**Relevant Coursework:** **Electrical Engineering:** Magnetic Resonance Imaging, Digital & Analog Signal Processing, Fields and Waves, Semiconductor Electronics, Digital Electronics, Digital Systems Lab **Computer Engineering/Science:** Machine Learning for Signals, Mobile Computing, Data Science, ML/AI, Data Structures and Algorithms, Computer Systems and Programming, **Physics:** Electricity and Magnetism, Quantum, Thermal, Mechanics, **Math:** Optimization, Probability Theory, Differential Equations, Linear Algebra, Calculus I-III

**Mountain Lakes High School, Mountain Lakes, NJ – Varsity Basketball Captain, ACT: 35/36 SAT Math 2 Subject test 800/800**

### Honors and Awards

**ECE Departmental Scholarship:** 2023-2024 recipient of Daniel W. and Carol A. Dobberpuhl Scholarship

**ECE Departmental Award:** 2022-2023 recipient of Ellery B. Paine Outstanding Junior Award

**ECE Departmental Scholarship:** 2022-2023 recipient of Bradley A. Simmons Memorial Scholarship

### Research Experience

**Senior Research project & Thesis**, Champaign, IL

**August 2023 – Ongoing**

Working under Dr. Minh Do as lead researcher exploring mmWave Radar for Person Identification and Classification

- Using Texas Instruments single chip radar alongside Intel's RealSense RGB-D camera to create/train a system to Identify human activity and perform person classification from micro-Doppler signatures in radar returns.

**Independent Study**, Champaign, IL

**January 2023 – May 2023**

Mentored by Dr. Venugopal Veeravalli as I explored Out-of-Distribution (OOD) detection in machine learning.

- Executed experiments employing a multiple-hypothesis testing framework with conformal p-values for various statistics.
- Tested Resnet and Densenet architectures on standard image datasets, assessing the effectiveness of Energy, Mahalanobis distance, and Gram Matrix statistics for OOD detection.

**Undergraduate Researcher**, Champaign, IL

**January 2022 – May 2022**

Researcher for Dr. Zhen Peng in running experiments on Reconfigurable Intelligent Surfaces

- Assembled RIS enhanced wireless communication testbed using Software Defined Radio transceivers and setup basic modulation transmission and reception using GNURadio

### Industry Experience

**Expedition Technologies, Machine Learning & Digital Signal processing Intern**, Herndon, VA

**May 2023 – August 2023**

Worked on a team designing a real-time system that could detect anomalous RF activity through combined DSP and ML

- Developed two automatic segmentation/labelling methods for RF spectrograms - one using Statistics/DSP to establish a noise ceiling then generate bounding boxes and another using a Neural Network with a U-Net backbone.
- Designed Flask & Vue.js based Data Labelling tool for RF data with automatic and manual labelling options as well as the Integrated ability to finetune new model checkpoints for improved labelling accuracy.
- Demonstrated the use of polyphase channelizers as a lower latency alternative to the spectrogram.

**BAE systems Electronic Systems, Electrical Engineering Intern**, Wayne, NJ

**May 2022 – August 2022**

Mentored by Radio Frequency engineers as I rotated through work in Design, Simulation, Analysis, and Testing

- Used Spectrum Analyzer, Network Analyzer, and Signal Generator to test for proper functionality of Radio Frequency Circuit boards involved in a Link-16 terminal often used in classified aircraft communication systems.
- Gained a macroscopic understanding of complex electrical and computer systems pertaining to Government Defense

**Colgate-Palmolive, Software Engineering Intern**, Piscataway, NJ

April 2021 – August 2021

Served as the Sole developer on an enterprise project to automate the creation of Google Groups based on a variety of criteria and built a Web App to allow employees to build their own automatically maintained groups.

- Performed Full stack web application development using a Node.JS backend, Vue.JS frontend, with GCP Data storage/interaction.
- Self-taught all skills used in a timely fashion while initially working and completing schoolwork/final exams in parallel.

## Teaching and Mentoring Experience

**Engineering Orientation TA**, Champaign, IL

January 2023-present

Instructor and mentor to a group of 21 freshman ECE students in an Engineering Orientation course, providing insights into the College of Engineering and ECE department.

- Conduct bi-weekly hour-long classroom sessions to deliver essential resources, share personal experiences, and prepare students for the academic rigor of ECE.
- Foster individualized support by conducting regular one-on-one meetings with each student, assessing how they are adapting, and addressing specific academic inquiries.

## Personal

**National Organization of Business and Engineering 2021 Project:** Text mood identifier using combined unigram and bigram model

**AI @ UIUC Social Director:** Served as social director for AI @ UIUC club. Fully organized social events throughout semester.

**Face Recognition System:** Used PCA and LDA alongside OpenCV to create a real time face recognition/tracking system.

**FPGA Projects:** Hardware lock picking video game with rotational motion and multiple levels.

**Skills:** NumPy, Scikit-Learn, PyTorch, Python, JavaScript, Electronics

**Security clearance:** I have a Secret Level Government Security Clearance.

**Hobbies:** Fitness, Piano, Wakeboarding, Investing.

## Academic History

676579195 Charles R. Plonski (Charlie)  
Sep 14, 2023 05:28 pm

 To print your University of Illinois academic history, follow the print instructions for your web browser as you would to print any web page. For example, with Internet Explorer, select the Print option from the File menu.

In the Degree Information section, you may see multiple sought degree records. To view your current degree, click on the View Student Information link at the bottom of the page.

 Note: Academic standing is reviewed by your college and is subject to change.

 This is NOT an Official Transcript.

[Transfer Credit](#) [Institution Credit](#) [Transcript Totals](#) [Courses in Progress](#)

### **Transcript Data**

#### **STUDENT INFORMATION**

**Name :** Charles R. Plonski

**Birth Date:** Mar 20, 2002

#### **Curriculum Information**

##### **Most Recent Program(s)**

**College:** Grainger Engineering

**Major and Department:** Electrical Engineering, Electrical & Computer Eng

**Minor:** Computer Science

\*\*\*This is NOT an Official Transcript\*\*\*

**TRANSFER CREDIT ACCEPTED BY INSTITUTION****-Top-**

FA20:		ACT/SAT						
Subject	Course	Title		Grade	Credit Hours	Quality Points		R
RHET	105	Writing and Research		PS	4.000			0.00
		Attempt Hours	Passed Hours	Earned Hours	GPA Hours	Quality Points	GPA	
<b>Current Term:</b>		0.000	4.000	4.000	0.000	0.00		0.00

⚠ This is NOT an Official Transcript.

**FA20:** Advanced Placement Tests

Subject	Course	Title		Grade	Credit Hours	Quality Points		R
CS	101	Intro Computing: Engrg & Sci		PS	3.000			0.00
CS	102	Little Bits to Big Ideas		PS	4.000			0.00
MATH	220	Calculus		PS	5.000			0.00
MATH	231	Calculus II		PS	3.000			0.00
PHYS	101	College Physics: Mech & Heat		PS	5.000			0.00
PHYS	2--	Test-Based Credit		PS	4.000			0.00
PHYS	211	University Physics: Mechanics		PS	4.000			0.00
RHET	105	Writing and Research		PS	4.000			0.00 E
STAT	100	Statistics		PS	3.000			0.00
		Attempt Hours	Passed Hours	Earned Hours	GPA Hours	Quality Points	GPA	
<b>Current Term:</b>		0.000	31.000	31.000	0.000	0.00		0.00

⚠ This is NOT an Official Transcript.

**SU21:** Oakton College

Subject	Course	Title		Grade	Credit Hours	Quality Points		R
CS	173	Discrete Structures		A	3.000			12.00
		Attempt Hours	Passed Hours	Earned Hours	GPA Hours	Quality Points	GPA	
<b>Current Term:</b>		3.000	3.000	3.000	3.000	12.00		4.00

⚠ This is NOT an Official Transcript.

**INSTITUTION CREDIT** **-Top-**

Term: Fall 2020 - Urbana-Champaign

College: Grainger Engineering

**Major:** Computer Engineering**Academic Standing:** Good Standing**Additional Standing:** Deans List

Subject	Course	Level	Title	Grade	Credit Hours	Quality Points	R
CHEM	102	1U	General Chemistry I	A	3.000	12.00	
CHEM	103	1U	General Chemistry Lab I	A	1.000	4.00	
ECE	120	1U	Introduction to Computing	A	4.000	16.00	
ECON	102	1U	Microeconomic Principles	A	3.000	12.00	
ENG	100	1U	Engineering Orientation	S	0.000	0.00	
MATH	241	1U	Calculus III	A+	4.000	16.00	
MUS	169	1U	Unit One Music Lessons	S	0.000	0.00	

**Term Totals (Undergrad - Urbana-Champaign)**

	Attempt Hours	Passed Hours	Earned Hours	GPA Hours	Quality Points	GPA
<b>Current Term:</b>	15.000	15.000	15.000	15.000	60.00	4.00
<b>Cumulative:</b>	15.000	15.000	15.000	15.000	60.00	4.00

⚠ This is NOT an Official Transcript.

**Term: Spring 2021 - Urbana-Champaign****College:** Grainger Engineering**Major:** Computer Engineering**Academic Standing:** Good Standing**Additional Standing:** Deans List

Subject	Course	Level	Title	Grade	Credit Hours	Quality Points	R
ECE	110	1U	Introduction to Electronics	A	3.000	12.00	
ECE	220	1U	Computer Systems & Programming	A	4.000	16.00	
KIN	249	1U	Sport & Modern Society	A+	3.000	12.00	
MATH	286	1U	Intro to Differential Eq Plus	A	4.000	16.00	
MUS	169	1U	Unit One Sem Instruct in Music	S	0.000	0.00	
PHYS	212	1U	University Physics: Elec & Mag	A+	4.000	16.00	

**Term Totals (Undergrad - Urbana-Champaign)**

	Attempt Hours	Passed Hours	Earned Hours	GPA Hours	Quality Points	GPA
<b>Current Term:</b>	18.000	18.000	18.000	18.000	72.00	4.00
<b>Cumulative:</b>	33.000	33.000	33.000	33.000	132.00	4.00

⚠ This is NOT an Official Transcript.

**Term: Fall 2021 - Urbana-Champaign**

**College:** Grainger Engineering  
**Major:** Computer Engineering  
**Academic Standing:** Good Standing  
**Additional Standing:** Deans List

Subject	Course	Level	Title	Grade	Credit Hours	Quality Points	R
CS	225	1U	Data Structures	A	4.000	16.00	
ECE	210	1U	Analog Signal Processing	A	4.000	16.00	
MUS	133	1U	Introduction to World Music	A+	3.000	12.00	
PHYS	213	1U	Univ Physics: Thermal Physics	A	2.000	8.00	
PHYS	214	1U	Univ Physics: Quantum Physics	A+	2.000	8.00	

**Term Totals (Undergrad - Urbana-Champaign)**

	Attempt Hours	Passed Hours	Earned Hours	GPA Hours	Quality Points	GPA
<b>Current Term:</b>	15.000	15.000	15.000	15.000	60.00	4.00
<b>Cumulative:</b>	48.000	48.000	48.000	48.000	192.00	4.00

⚠ This is NOT an Official Transcript.

**Term: Spring 2022 - Urbana-Champaign**

**College:** Grainger Engineering  
**Major:** Electrical Engineering  
**Academic Standing:** Good Standing  
**Additional Standing:** Deans List

Subject	Course	Level	Title	Grade	Credit Hours	Quality Points	R
ECE	313	1U	Probability with Engrg Applic	A	3.000	12.00	
ECE	314	1U	Probability in Engineering Lab	A	1.000	4.00	
ECE	329	1U	Fields and Waves I	A+	3.000	12.00	
ECE	385	1U	Digital Systems Laboratory	A	3.000	12.00	
MUS	132	1U	Popular Music Studies	A	3.000	12.00	

**Term Totals (Undergrad - Urbana-Champaign)**

	Attempt Hours	Passed Hours	Earned Hours	GPA Hours	Quality Points	GPA
<b>Current Term:</b>	13.000	13.000	13.000	13.000	52.00	4.00
<b>Cumulative:</b>	61.000	61.000	61.000	61.000	244.00	4.00

⚠ This is NOT an Official Transcript.

**Term: Fall 2022 - Urbana-Champaign**

**College:** Grainger Engineering

**Major:** Electrical Engineering**Academic Standing:** Good Standing**Additional Standing:** Deans List

Subject	Course	Level	Title	Grade	Credit Hours	Quality Points	R
DANC	100	1U	Intro to Contemporary Dance	A+	3.000	12.00	
ECE	310	1U	Digital Signal Processing	A+	3.000	12.00	
ECE	350	1U	Fields and Waves II	A+	3.000	12.00	
ECE	365	1U	Data Science and Engineering	A+	3.000	12.00	
ECE	448	1U	Artificial Intelligence	A	3.000	12.00	

**Term Totals (Undergrad - Urbana-Champaign)**

	Attempt Hours	Passed Hours	Earned Hours	GPA Hours	Quality Points	GPA
<b>Current Term:</b>	15.000	15.000	15.000	15.000	60.00	4.00
<b>Cumulative:</b>	76.000	76.000	76.000	76.000	304.00	4.00

⚠ This is NOT an Official Transcript.

**Term: Spring 2023 - Urbana-Champaign****College:** Grainger Engineering**Major:** Electrical Engineering**Academic Standing:** Good Standing**Additional Standing:** Deans List

Subject	Course	Level	Title	Grade	Credit Hours	Quality Points	R
ECE	396	1U	Honors Project	A+H	2.000	8.00	
ECE	420	1U	Embedded DSP Laboratory	A+	2.000	8.00	
ECE	434	1U	Mobile Computing & Application	A	4.000	16.00	
ECE	490	1U	Introduction to Optimization	A	3.000	12.00	
ENG	377	1U	ELA Leadership Training	A	1.000	4.00	
PSYC	230	1U	Perception & Sensory Processes	A+	3.000	12.00	

**Term Totals (Undergrad - Urbana-Champaign)**

	Attempt Hours	Passed Hours	Earned Hours	GPA Hours	Quality Points	GPA
<b>Current Term:</b>	15.000	15.000	15.000	15.000	60.00	4.00
<b>Cumulative:</b>	91.000	91.000	91.000	91.000	364.00	4.00

⚠ This is NOT an Official Transcript.

**TRANSCRIPT TOTALS (UNDERGRAD - URBANA-CHAMPAIGN) -Top-**

	<b>Attempt Hours</b>	<b>Passed Hours</b>	<b>Earned Hours</b>	<b>GPA Hours</b>	<b>Quality Points</b>	<b>GPA</b>
<b>Total Institution:</b>	91.000	91.000	91.000	91.000	364.00	4.00
<b>Total Transfer:</b>	3.000	38.000	38.000	3.000	12.00	4.00
<b>Overall:</b>	94.000	129.000	129.000	94.000	376.00	4.00

⚠ This is NOT an Official Transcript.

### COURSES IN PROGRESS -Top-

Term: Fall 2023 - Urbana-Champaign

College: Grainger Engineering  
Major: Electrical Engineering

<b>Subject</b>	<b>Course</b>	<b>Level</b>	<b>Title</b>	<b>Credit Hours</b>
CS	545	1U	Machine Learning for Signals	4.000
ECE	340	1U	Semiconductor Electronics	3.000
ECE	480	1U	Magnetic Resonance Imaging	3.000
ECE	496	1U	Senior Research Project	2.000

⚠ This is NOT an Official Transcript.

**RELEASE: 8.7.1**

© 2023 Ellucian Company L.P. and its affiliates.

# GIDEON POPOOLA

## SKILLS

SQL · Python · TensorFlow · Keras · R ·  
Apache Spark · Big data · Hadoop/Hive ·  
Machine Learning Algorithms ·  
Hypothesis Testing · NLP · Scala · AWS ·  
Pytorch · PowerBi · Data Visualization

## STRENGTHS

### Leadership

Managed a 7-member cross-functional (data analysts, data engineers, and sales) team and coordinated with five business partners toward the successful launch of an automated data reporting platform.

### Communication

Led a remote team of 5+ developers in assessing project outcomes and working closely with senior leadership to prioritize future features.

## PROJECTS

### Statistical Model for Disease Prediction.

- Developed a mixed-effect negative binomial model to predict patients' risk of developing kidney diseases, improving kidney disease diagnosis by 20%.

### Customer Churning Prediction

- Designed and implemented an XGBoost algorithm to predict telecommunications customer churn, resulting in a remarkable 65% increase in customer retention.

### Machine Learning Model

- Pioneered a cutting-edge deep learning-based feature engineering technique to effectively tackle class imbalance problems, leading to substantial improvements in key metrics such as a 5% increase in AUC over previous algorithms.

## Data Scientist

📞 406-219-6160 @ popolagideonolasunbo@yahoo.com  
🔗 <https://www.linkedin.com/in/gideon-olasunbo-popoola/>📍 Bozeman, MT

## EDUCATION

### Masters in Computer Science

01/2022 - Present

**Montana State University**

### Bachelors of Computer Science

11/2015 - 08/2019

**Olabisi Onabanjo University**

## EXPERIENCE

### Graduate Assistant

01/2022 - Present

**Montana State University**

Bozeman, MT

- Designed and implemented interactive dashboards using SQL and Power BI, showcasing 30+ KPIs that provided insights into various departmental operations.
- Led end-to-end data collection for 10+ research projects, employing a blend of qualitative and quantitative methods.
- Applied parametric and non-parametric statistical tests in R Studio to evaluate the effect of clinical treatments on patients, leading to a significant 45% reduction in clinical decision-making time.
- Led collaboration with 30+ data analysts, data scientists, and project managers on concurrent projects in Q1 2022, resulting in 3+ publications.

### Data Scientist

09/2019 - 12/2021

**Africa4Ai**

Pretoria , SA

- Developed ten KPI dashboards focusing on infection control and medication administration processes, enhancing staff's ability to track patient progress and leading to an impressive 87% improvement in monitoring efficiency.
- Designed and implemented mixed-effect logistic regression models to forecast patient demand and optimize resource allocation, improving resource allocation by 16%.
- Created data visualization dashboards with Python, Tableau, and SQL, enhancing data accessibility and interpretation. Resulted in a 49% boost in decision-making efficiency.
- Constructed an interactive dashboard for 150+ private sector clients, facilitating exploration of employment and investment opportunities across 30+ geographies.
- Engineered a personalized machine learning algorithm for 750+ clients, revolutionizing customer retention and driving exceptional monthly sales growth of 65% by leveraging complex data analysis and implementing tailored strategies based on individual customer preferences and behaviors.
- Developed an NLP system employing 11+ natural language processing techniques to automatically classify 5.5k emails as either spam or advertising, enhancing email filtering and user experience.

### Data Analyst Intern

04/2018 - 09/2018

**Olabisi Onabanjo University**

Ogun, NG

- Compared different supervised machine learning algorithms to optimize the prediction of over 100k+ student academic performance, reducing dropout rate from 17% to 9%.

- Spearheaded a comprehensive template that streamlined the billing process and helped slash annual overhead costs by 39%.

- Created explanatory models of use cases for presentations to 100+ project stakeholders and 40+ non-technical audiences.

- Designed and implemented an intelligent solution that reduced manual data entries from 89% to 5%, decreasing processing time from 20+ days to 48 minutes.

# ABDUL MATIN

## GRADUATE INTERN

### CONTACT

-  970-694-8150
-  abdul.matin@colostate.edu
-  linkedin.com/in/abdulmatinmuaz
-  github.com/ammuaj
-  G. Scholar: bit.ly/3lb0vkm
-  Fort Collins, CO, USA

### SKILLS

- Python
- React
- JavaScript
- TypeScript
- MongoDB
- Git
- SQL
- Java, Scala
- Map Reduce
- Hadoop
- Apache Spark
- ANN, CNN, GNN, VGG, GAN
- ResNet, MobileNet, CapsuleNet
- PyTorch, TensorFlow

### EDUCATION

#### PhD in Computer Science

#### Colorado State University

Graduation: Dec 2025 (Expected)

GPA: 4.0

#### M.Sc. in Computer Science

#### Colorado State University

Graduation: Dec 2022

GPA: 3.79

#### B.Sc. in CS & Engineering

#### RUET, Bangladesh

Graduation: Jan 2016

GPA: 3.89 (Topper)

### CORE COMPETENCIES

Innovative Thinking | Teamwork | Active Learning | Self-Awareness | Communication | Analytical Capability

### WORK EXPERIENCE

#### ML Engineer and Student Researcher

Big Data Lab, Colorado State University Aug 2022-Present

- Collaborated with 10 team members to develop a REACT application for an interactive large-scale visualization support system for different large-scale geospatial data
- Conducted in-depth, groundbreaking research in the realm of big data analytics through the utilization of cutting-edge deep neural networks.
- Publications: IEEE BigData 2023 [accepted], ICLR 2024 [submitted], TransAI 2022, eScience 2022

#### Software Engineer Intern

Big Data Lab, Colorado State University May 2022 - Aug 2022

- Led and managed a cross-functional team of 15 members to build a powerful React application to support revolutionized large-scale data analytics
- Improved the query processing speed and enabled real-time insights for informed decision-making
- Developed an intuitive user interface using JavaScript and React, resulting in improved page load time
- Incorporated Python for the data ingestion, QA/QC, and the large-scale data analytical backend
- Employed the query optimization and data retrieval from MongoDB for geospatial query support backend in Java to reduce the server response time by 20%

#### Lecturer and Mentor

RUET, Rajshahi, Bangladesh Oct 2016 - Dec 2020

- Designed and delivered a comprehensive curriculum on OOP, Data Structures, and algorithms to 3 classes of 60+ students; and received outstanding feedback from 90% of students
- Engaged in rigorous academic research in the cutting-edge machine-learning field
- Took the lead as a mentor, guiding and co-ordinating 12 students to pursue knowledge and research at the undergraduate level, resulting in more than 15 prestigious publications

### PROJECTS

#### Python | TensorFlow | CNN | Teamwork

Aug 2021 - Dec 2021

An application to predict wildfire from weather data by analyzing historical data

- Conducted the team project as a part of the requirements of the CS555 course at CSU
- Implemented ML model using 1DCNN in Python to predict wildfire based on past weather data

#### Java | Teamwork

Aug 2021 - Dec 2021

A Distributed, Replicated, And Fault-Tolerant P2P File Management System

- Employed scalable and parallel peer-to-peer concurrent file accessing system with 3-stage replica in Java as part of the team project requirement of CS555 course at CSU

#### Python | TensorFlow| ANN

Jan 2021 - May 2021

An Efficient pre-owned Car Price Prediction System using artificial neural network (ANN)

- Developed a more accurate machine learning based solution to predict the pre-owned car price

# ADERINSOLA OSINOWO

678-777-1319 | [aderinsolaosinowo@gmail.com](mailto:aderinsolaosinowo@gmail.com) | <https://www.linkedin.com/in/aderinsolaosinowo>

## EDUCATION

**Georgia State University Honors College** | Cumulative GPA: 3.77 Atlanta, GA  
Bachelor of Interdisciplinary Studies: Biomedical Science & Enterprises | Minor: Computer Information Systems. December 2025  
**Awards:** Dean's List Academic Achievement 2021-2023, HOPE Scholarship Recipient, WomenLead Programmatic Scholarship

## PROFESSIONAL EXPERIENCE

**PwC** | *Non-Profit Consulting Extern* | Remote July 2023- September 2023  
• Engaged in a ParagonOne externship with PwC and Breakthrough Collaborative, elevating their social media presence.  
• Assessed, improved, and presented data-driven recommendations with interactive visuals to PwC and BTC executives.  
• Identified influencers and proposed a 65% engagement increase through Tik-Tok and Instagram.  
• Recognized with a certificate of Successful Completion for impactful contributions during the externship.

**Code Wiz** | *Coding & Robotics Mentor* | Lawrenceville, GA May 2023- August 2023  
• Guided 10+ students daily in creating robots using LEGO® and Scratch Programming for coding.  
• Delivered game development instruction to campers using Roblox Studio, nurturing their creativity and technical abilities.  
• Instructed students of various age groups in elementary coding techniques.  
• Collaborated with coding mentors to stay current with technology and software advancements.

**Georgia State University** | *Student Assistant, International Student & Scholar Services (ISSS)* | Atlanta, GA September 2022-Present  
• Efficiently manages the distribution of over 25 confidential USCIS reports weekly to students and scholars.  
• Streamline and optimizes international student event logistics under ISSS supervision for enhanced attendee experiences.  
• Provides expert counsel to 35+ international students weekly, addressing F-1, J-1, and H-1B Visa Status inquiries  
• Serves as a Call Center Representative, handling 50+ weekly calls and guiding students to International Advisors as needed.

**Kids R Kids Learning Academy** | *Assistant Teacher* | Lawrenceville, GA June 2021-July 2022  
• Instructed students aged 5 to 10 in core-curriculum subjects, encompassing lower-level and upper-level Mathematics.  
• Trained four assistant teachers in the Kids-R-Kids Teaching Curriculum and Georgia's Childcare Safety protocols.  
• Digitized and managed crucial data for 60+ students, resulting in a 55% faster retrieval of files for advisors.  
• Cultivated a nurturing classroom environment, fostering students' socio-emotional development and academic growth.

## LEADERSHIP & PROFESSIONAL DEVELOPMENT

**Prestiged Org Sisterhood** | *Vice President* | Atlanta, GA May 2023-Present  
• Orchestrate successful events, ensuring high attendance and seamless planning within the organization.  
• Deliver presentations to 25 members, sparking discussion-based conversations and fostering networking opportunities.  
• Cultivate a supportive community, empowering members to maximize university resources effectively.

**Project: FLOW** | *Co-Service Chair* | Atlanta, GA February 2023-Present  
• Collaborate with team members to assemble and distribute 45+ period packages to shelters in Downtown Atlanta.  
• Spearhead the procurement of community service resources, enabling Project: Flow to expand its assistance to those in need.  
• Assist in period-product restocking for university restrooms, ensuring students have access to essential hygiene products.

**Serve @ State** | *Director of Programming* | Atlanta, GA January 2023-Present  
• Organize bi-weekly events to drive civic engagement and community outreach goals.  
• Partner with other student organizations at GSU to create volunteer opportunities through collaborative events.  
• Facilitate the coordination of impactful events to support the community.

**GSU Panther LEAP Leadership Program** | *LEAPer* | Atlanta, GA September 2022-December 2022  
• Chosen for a four-month leadership program centered on promoting social justice, inclusion, and civic engagement at GSU.  
• Crafted and showcased a 3-D diorama detailing the End SARS movement in Nigeria, addressing police brutality concerns to LEAP mentors.  
• Collaborated with local non-profit Lost-n-Found Youth to organize and manage thrift store clothing inventory.

## AFFILIATIONS & SKILLS

**Affiliation(s):** GSU Women LEAD, GSU Consulting Club, GSU Honors Program

**Language:** Yoruba (Professional Working Proficiency)

**Skills:** Pediatric CPR & First Aid Training, Data Management & Analysis, Presentation Skills, Executive Presence

**Software:** Proficient in MS Office Suite, Sunapsis, Codio