

EDUCATION

University of Illinois Urbana-Champaign - **MS, Statistics (GPA: 4/4)** **08/2023 — Ongoing**
Coursework: Time Series Analysis, Advanced Regression Analysis, Biostatistics, Mathematical Statistics.

Manipal Institute of Technology - **BTech, Computer Science & Engineering (GPA: 4/4)** **07/2018 — 07/2022**
Dean's List for procuring a GPA of **10/10** in semester VIII, VII, **9.85/10** in semester VI and **9.55/10** in semester II.

EXPERIENCE

Graduate Research Assistant, [University of Illinois Urbana-Champaign](#) **01/2024 — Present**

Advisor: Prof. Sihai Dave Zhao

- Developing advanced computational models to process and analyze large-scale spatial transcriptomics datasets, leveraging deep learning techniques for enhanced accuracy and efficiency.
- Designed and implemented novel diffusion models to optimize nearest neighbor search algorithms.
- Developed KDE++, a data structure tailored for efficient KDE in extremely large datasets, facilitating faster and accurate data analysis. Demonstrated a computational speedup of over 45x by integrating KDE++ with parallel processing and data handling strategies.
- Collaborating with cross-functional research teams (with Prof. Han Hee Sun, Department of Chemistry, UIUC) to integrate computational models into broader biological research initiatives, enhancing the interpretability and applicability of spatial transcriptomics data.
- Conducting comprehensive evaluations and benchmarks to validate model performance and robustness, ensuring reliability for large-scale data modeling applications.

Project in progress

Machine Learning Intern, [Bayer Research & Development](#) **05/2024 — Present**

Advisor: Dr. Jiarui Li

- Research and development of rank correlation analyses to determine optimal moisture levels for maximizing data quality and yield predictions.
- Developing advanced models to evaluate the strategic value of aggressive harvesting practices, focusing on their impact on yield rank and resource allocation.

Project in progress

Independent Research Associate, [Samsung Research India](#) **06/2022 — 12/2022**

Mentor: Pushkar Raj Singh

- Researched deep learning and geometrical methods for near-infrared images.
- To combat lack of data, developed t-SNE and Uniform Manifold Approximation and Projection models for dimensionality reduction and predictive analysis.
- Formulated and authored a novel architecture employing auto-encoders to project high dimensional data from the NIR images into a low dimensional manifold.
- Applied a Gaussian Mixture Model to evaluate new, unfamiliar data patterns. *Paper: [1]*.

Software Development Engineer, [Citrix Research and Development](#) **07/2022 — 07/2023**

- Engaged in research, development and testing for the Citrix Workspace App for HTML5 & ChromeOS. Helped to drive the expansion of the client base from **700K to 1M monthly active users**.
- Optimized client-side printing algorithms, reducing delays by 20%, resulting in the delivery of **Q1 2023's best feature**.
- Analyzed user data to fix bugs and automated cloud processes affecting over **100K users**, ensuring regulatory compliance.

Software Development Intern, [Citrix Research and Development](#) **01/2022 — 06/2022**

- Gained in-depth proficiency with Google Analytics 4, understanding the potential to optimize data visualization and analytics. Created comprehensive documentation used by over 10+ teams across CWA to integrate GA4.
- Developed ethics and time-sensitive features affecting over **100,000 users** in the EU region.

- Devised a quantitative model to analyse trade roots to reduce cost of fuel for fleet management.
- Worked on analysing and utilising the data to track drivers for fleet management and score drivers based on over-speeding, fatigued driving, sharp turns, SLA violations and other metrics.
- Wrote unit tests for the behavioural system to test the functionalities of the source code.

- Established the first biology based student project in Manipal and obtained **\$17K funding** by creating proposals.
- Directing lab experiments based on thorough data analysis to model a probiotic to mitigate affects of methyl mercury poisoning.
- Surveyed awareness of MMP and GMOs; modeled acceptance of GMO-containing prebiotics.
- Awarded **Gold medal** for integrating human practices in solution.

SELECTED PUBLICATIONS

Journal and Conference Submissions

1. **Navya Gupta**, Pushkar Raj Singh. *Assessment of Near Infrared Images through Dimensionality Reduction*
2. **Navya Gupta**, Gokul Puthumanaillam. *You Only Look For Context(with a little help from rules): Imitation Learning based context aware attention module.*

TECHNICAL SUMMARY

Languages & Tools	Git, C / C++, Java, SQL, Python, HTML, CSS, JavaScript, \LaTeX , MATLAB, Linux toolchain, GNU Octave, R, PowerBI.
Libraries and Frameworks	pandas, NumPy, SciPy, scikit-learn, Keras, LibROSA, Pillow, OpenCV TensorFlow, PyTorch, fastai, PyStan, matplotlib, seaborn.

SELECTED PROJECTS (VIEW ALL PROJECTS AT: [GITHUB.COM/WONKYVAMP](https://github.com/WonkyVamp))

Biomarker Discovery in Cancer Gene

github.com/WonkyVamp/geneSelection

- Controlled the false discovery rate in high-dimensional gene expression data for identifying potential cancer biomarkers for breast, lung, colorectal, thyroid, and ovarian cancers.
- Conducted feature selection using regression and decision trees with and without a knockoff filter and gradient boosting and logistic regression as the classification model achieving an **accuracy of 99.97%**.

ProteinPalette

github.com/WonkyVamp/ProteinPalette

- Analyzed protein content in chemicals using predictive analysis and spectroscopy.
- Compiled 1000+ bands dataset with mobile and benchmark sensors for spectroscopic absorbance.
- Preprocessed data with scatter correction and SNV; applied convolution and Savitzky-Golay filtering.
- Achieved **R2 value of 0.91** using a deep belief statistical model with 300 wetlab samples.

TextQuratton

github.com/WonkyVamp/TextQuratton

- Developed TextQuratton, an app to parse and detect fraudulent handwritten bills.
- Introduced a context-aware method with a rule-based learning model for dynamic text tagging.
- Utilized spatial transformers, CRNN, and bidirectional-LSTM for bill parsing, achieving **93.67% accuracy**.
- Secured **runners-up position at HackRx hackathon**.

HONORS AND AWARDS

- **1st Runner up, Bajaj HackRx** amongst 500+ teams in the nation-wide hackathon.
- **Department topper**, Computer Science and Engineering Department for the year 2021-2022.
- **Excellence in academics award** for GPA of 10 in semester VIII & VII, 9.85 in semester VI & 9.55 in semester II.
- **Gold Medal, iGEM 2020** for Integrated Human Practices, Model and Science Communication.

TEACHING AND MENTORSHIP

- **Teaching Assistant (ASTR-121 FA23, SP24)**, responsible for leading discussions, grading and holding office hours.
- **Teach Code for Good**, 120+ hours invested in teaching underprivileged children.
- **Manipal BioMachines**, founding member and head of web development.