Sambit Panda

Cary, NC 27513 | **US Citizen**

919-637-6272 | sampanda501@gmail.com | linkedin.com/in/sampan501 | github.com/sampan501 | sampan.me

SUMMARY

- Highly motivated professional with 10+ years of research experience; interests include generative AI, machine learning, data science, statistics, and healthcare
- Author of 14 publications (h-index: 7, 200+ citations)
- 7+ years of experience using Python and R to develop data science solutions in academic and industry settings

EDUCATION

Johns Hopkins Medical Institute   **Baltimore, MD**
PhD, Biomedical Engineering   *Jul 2020 – Dec 2024*
- Advisor: Joshua T. Vogelstein
- Received the NIH T32GM119998 grant

Johns Hopkins University   **Baltimore, MD**
MSE, Biomedical Engineering   *Aug 2018 – May 2020*
- Advisor: Joshua T. Vogelstein

NC State University & UNC Chapel Hill   **Raleigh & Chapel Hill, NC**
BS, Biomedical Engineering & Biology   *Aug 2014 – May 2018*
- Advisor: Leslie Sombers
- Received the Goodnight Scholarship (full ride) and the National Merit Corporate Scholarship.

WORK EXPERIENCE

MATRIX AI Consortium   **Dec 2024 – Present**
AI Research Scientist   *Remote*
- Built an AI agent to aid emergency physicians with decision-making using a **Python** (via **FastAPI**, **LangChain)**, and **OpenAI/Gemini API** backend and a **React** frontend (via **Vercel AI SDK**, **Next.js**, and **Tailwind CSS**)
- Collaborated with leading trauma care physicians in Texas and Eric Horvitz (CSO at Microsoft) to build a base of knowledge of the assistant and quantify decision uncertainty of provided recommendations
- Devised a framework for evaluating LLMs (crafting and distributing a survey to ~50 clinicians in the process) and collaborated with top AI researchers to place performance guarantees on LLMs
- Leveraged ~400000 patients' historic data and new data collected by 7 trauma centers across Texas (using **SQL** with data stored in **Google BigQuery**) to develop an AI-driven geospatial tool to inform trauma care policy using **Python** (via **Dash** and **PyTorch**)
- Held a workshop attended by 50+ participants on how to use Generative AI for Biomedical decisions

NeuroData Lab, Johns Hopkins   **Jan 2019 – Dec 2024**
Researcher   *Baltimore, MD*
- Developed multiple algorithms, notably KMERF (random forest-based hypothesis test), Nonparametric MANOVA (a nonparametric multivariate k-sample test), Fast Dcorr (fast approximation to the distance correlation test), and Causal Dcorr (distance correlation for causal inference)
- Authored 11 publications (5 first author, ~150 citations) related to early cancer detection, random forest, neural networks, causal inference, and hypothesis testing using **Python** packages like **TensorFlow** and **PyTorch**
- Created and maintained open-source **Python** packages like hyppo (~150 users, 200+ stars, ~100 forks) and treeple (50+ stars, ~20 forks); ported algorithms from these packages into SciPy.
- Developed and tested code using **Git**, **Docker**, Cloud Services (**AWS EC2/S3**, **Azure VM**), Continuous Integration (**CircleCI**, **Travis CI**, etc.), and **Python** packages (**pandas**, **scikit-learn**, etc.)
- Collaborated with Bert Vogelstein, a renowned scientist in cancer genomics, on the MIGHT algorithm that quantifies predictive information in liquid biopsy feature sets; used **Python** packages (treeple, **scikit-learn**, **pandas**, etc.); wrote manuscript in preparation for PNAS
- Served as SciPy symposium conference chair and reviewer; journal reviewer for SoftwareX; presented work at top conferences like the BRAIN PI meeting and GYSS
- Worked on a project annotating whole body CT scans using **Python, Unix shell scripts**

National Institutes of Environmental Health Sciences   **May 2023 – Jul 2023**

Data Scientist   *RTP, NC*
- Applied the KMERF algorithm (which I created) to discover relationships in neurological data using **Python** packages (**pandas**, **scikit-learn**, etc.) and **R**; won 1st place in poster competition
- Collaborated with researchers to publish two manuscripts: (1) neurotransmitter signaling from fear response in mice and (2) the development of a fiber photometry **R** package; developed tutorials interfacing **Python** and **MySQL**

Sombers Lab, NC State   Jan 2015 – May 2018
Research Assistant   *Raleigh, NC*
- Created a hydrogen-peroxide specific electrochemical sensor; published in ACS Analytical Chemistry.
- Analyzed data and engineered solutions for numerous additional projects using **MATLAB**.
- Presented research at top conferences like society of neuroscience (SfN) and Pittcon.

Burleson Research Technologies   May 2015 – Sep 2015
Intern   *RTP, NC*
- Tested pharmaceutical drugs on rats and mice through various methods such as oral gavage, *i.p.*, and *i.v.*

National Institutes of Environmental Health Sciences   Jun 2013 – Jan 2014
Research Assistant   *RTP, NC*
- Trained in several basic genetics and neuroscience techniques such as PCR, gel electrophoresis, etc.

PUBLICATIONS

Journal Articles
1. **Panda, S.***, Shen, C.*, Perry, R., Zorn, J., Lutz, A., Priebe, C. E., & Vogelstein, J. T. (2025). Universally Consistent K-Sample Tests via Dependence Measures. *Statistics and Probability Letters*, *216*(1), 110278. https://doi.org/10.1016/j.spl.2024.110278
2. Bridge, M. F., Wilson, L. R., **Panda, S.**, Stevanovic, K. D., Letsinger, A. C., McBride, S., & Cushman, J. D. (2024). FiPhA: An open-source platform for fiber photometry analysis. *Neurophotonics*, *11*(1), 014305. https://doi.org/10.1117/1.NPh.11.1.014305
3. Wilson, L. R.*, Plummer, N. W.*, Evsyukova, I. Y., Patino, D., Stewart, C. L., Smith, K. G., Konrad, K. S., Fry, S. A., Deal, A. L., Kilonzo, V. W., **Panda, S.**, Sciolino, N. R., Cushman, J. D., & Jensen, P. (2024). Partial or Complete Loss of Norepinephrine Differentially Alters Contextual Fear and Catecholamine Release Dynamics in Hippocampal CA1. *Biological Psychiatry Global Open Science*, *4*(1), 51–60. https://doi.org/10.1016/j.bpsgos.2023.10.001
4. Shen, C., **Panda, S.**, & Vogelstein, J. T. (2022). The Chi-Square Test of Distance Correlation. *Journal of Computational and Graphical Statistics*, *31*(1), 254–262. https://doi.org/10.1080/10618600.2021.1938585
5. Wilson, L. R., **Panda, S.**, Schmidt, A. C., & Sombers, L. A. (2018). Selective and Mechanically Robust Sensors for Electrochemical Measurements of Real-Time Hydrogen Peroxide Dynamics in Vivo. *Analytical Chemistry*, *90*(1), 888–895. https://doi.org/10.1021/acs.analchem.7b03770

Preprints
1. Curtis, S.*, **Panda, S.***, Li, A.*, Xu, H., Bai, Y., Ogihara, I., O'Reilly, E., Wang, Y., Dobbyn, L., Popoli, M., Ptak, J., Nehme, N., Silliman, N., Tie, J., Gibbs, P., Ho-Pham, L., Tran, B., Tran, T., Nguyen, T., Goggins, M., Wolfgang, C., Wang, T., Shih, I., Fader, A., Lennon, A. M., Hruban, R., Bettegowda, C., Gilbert, L., Kinzler, K., Papadopoulous, N., Vogelstein, B., Vogelstein, J. T.^, Douville, C.^ (2024). *Detecting and Combining Useful Sets of Predictive Variables*. Manuscript in preparation for PNAS.
2. **Panda, S.**, Palaniappan, S., Xiong, J., Bridgeford, E. W., Mehta, R., Shen, C., & Vogelstein, J. T. (2024). *hyppo: A Multivariate Hypothesis Testing Python Package*. Manuscript under review at JMLR. https://doi.org/10.48550/arXiv.1907.02088
3. **Panda, S.***, Shen, C.*, & Vogelstein, J. T. (2024). *Learning Interpretable Characteristic Kernels via Decision Forests*. Manuscript in preparation for ICML 2025. https://doi.org/10.48550/arXiv.1812.00029
4. Konishcheva, K., Leventhal, B., Koyama, M., **Panda, S.**, Vogelstein, J. T., Milham, M., Lindner, A.*, & Klein, A.* (2024). *Accurate and efficient data-driven psychiatric assessment using machine learning*. Manuscript under review at JCPP Advances. https://doi.org/10.31234/osf.io/sekfw
5. Bridgeford, E. W., Powell, M., Kiar, G., Noble, S., Chung, J., **Panda, S.**, Lawrence, R., Xu, T., Milham, M., Caffo, B., & Vogelstein, J. T. (2024). *When no answer is better than a wrong answer: a causal perspective on batch effects*. Manuscript under review at Imaging Neuroscience. https://doi.org/10.1101/2021.09.03.458920
6. Bridgeford, E. W., Chung, J., Gilbert, B., **Panda, S.**, Li, A., Shen, C., Badea, A., Caffo, B., & Vogelstein, J. T. (2024). *Learning sources of variability from high-dimensional observational studies*. Manuscript in preparation for Biometrika. https://doi.org/10.48550/arXiv.2307.13868

7. Xu, H., Dey, J., **Panda, S.**, & Vogelstein, J. T. (2022). *Simplest Streaming Trees.* arXiv. https://doi.org/10.48550/arXiv.2110.08483

8. Xu, H., Kinfu, K. A., LeVine, W., **Panda, S.**, Dey, J., Ainsworth, M., Peng, Y.-C., Kusmanov, M., Engert, F., White, C. M., Vogelstein, J. T., & Priebe, C. E. (2021). *When are Deep Networks really better than Decision Forests at small sample sizes, and how?.* arXiv. https://doi.org/10.48550/arXiv.2108.13637

Other Publications

1. **Panda, S.** (2020). *Multivariate Independence and k-sample Testing* [Thesis, Johns Hopkins University]. https://jscholarship.library.jhu.edu/handle/1774.2/62706

PROJECTS

iRemedyACT | *Python (LangChain, OpenAI & Gemini API, PyTorch), SQL, Next.js, Google Cloud*   **2024 – Present**
- A LLM agent to aid emergency physicians decision making at the point of care.
- A real-time geospatial model leveraging AI to give provide data-driven decisions for policy makers.
- Role: Creator and maintainer of both applications.

scipy.stats.multiscale_graphcorr | *Python, Cython*   **2019 – Present**
- Multiscale Graph Correlation is a powerful multivariate test (the 1st and only multivariate test in SciPy).
- Role: Ported this algorithm from hyppo and maintainer.

hyppo (originally mgcpy) | *Python (scikit-learn, pandas), CircleCI, Cloud (AWS, Azure)*   **2018 – Present**
- The first Python package for multivariate hypothesis testing, closing the gap with R (~150 users, 200+ stars, ~100 forks).
- Role: Creator and maintainer of this package.

treeple (originally scikit-tree) | *Python (scikit-learn, pandas), Cython*   **2023 – 2024**
- Extends scikit-learn decision trees to do oblique splits, manifold learning, hypothesis testing, etc. (50+ stars, ~20 forks).
- Role: Core contributor and maintainer of this package.

FiPhA | *R, Shiny*   **2023**
- A robust and user-friendly package for fiber photometry analysis.
- Role: Open-sourced this package and helped maintain it.

PRESENTATIONS

Talks

1. **Panda, S.**, & Cruz, C. (2025, May). *Generative AI for Biomedical Decisions* [Oral Presentation]. MATCH DICB AIM-AHEAD program, Virtual.

2. **Panda, S.**, Shen, C., Perry, R., Zorn, J., Lutz, A., Priebe, C. E., & Vogelstein, J. T. (2022, January). *Nonparametric MANOVA via Independence Testing* [Oral Presentation]. Global Young Scientists Summit, Virtual. https://www.youtube.com/watch?v=rJyuTwkgfjQ

3. **Panda, S.**, Wilson, L. R., & Sombers, L. A. (2018, May). *Hydrogen Peroxide, Dopamine, and Serotonin: Overlapping Chemical Systems Contribute to the Control of Dyskinetic Movements in the Rat During Chronic L-DOPA Treatment for Parkinson's Disease* [Capstone Presentation], Raleigh, NC, USA.

    → ***Won the Richard L. Blanton Outstanding Capstone Award for best capstone***

4. **Panda, S.**, Riley, S., Wiggins, K., Kathard, R., Alredge, T., & Krause, E. (2018, May). *Developing Solutions for Hand Spasticity* [Pitch] i4 Competition, RTP, NC, USA.

5. **Panda, S.**, Riley, S., Wiggins, K., Kathard, R., Alredge, T., & Krause, E. (2018, February). *Developing Solutions for Hand Spasticity* [Pitch] i4 Competition, RTP, NC, USA.

    → ***Won 1st place for pitch***

6. **Panda, S.**, Riley, S., Wiggins, K., Kathard, R., Alredge, T., & Krause, E. (2017, November). *Developing Solutions for Hand Spasticity* [Pitch] i4 Competition, RTP, NC, USA.

    → ***Won 1st place for pitch***

7. **Panda, S.**, & Lucas, S. (2016, February). *Surgical Site Infection Prevention* [Pitch] i4 Competition, RTP, NC, USA.

8. **Panda, S.**, & Lucas, S. (2015, October). *Surgical Site Infection Prevention* [Pitch] i4 Competition, RTP, NC, USA.

    → ***Won 1st place for pitch***

Posters

1. **Panda, S.**, Wilson, L. R., Stallone, J., Kendricks, D., Stevanovic, K., & Cushman, J. D. (2023, July). *Elucidating Relationships within Neurological Screening Batteries via Random Forest-Based Hypothesis Testing* [Poster Presentation] RTP, NC,

USA.

 → *Won 1st place for best Graduate poster*

2. **Panda, S.**, Shen, C., Perry, R., Zorn, J., Lutz, A., Priebe, C. E., & Vogelstein, J. T. (2021, June). *Nonparametric MANOVA via Independence Testing* [Poster Presentation] BRAIN Initiative Meeting, Virtual.

3. **Panda, S.**, Wilson, L. R., Schmidt, A. C., & Sombers, L. A. (2018, May). *Highly Selective and Mechanically Robust Sensors for Electrochemical Measurements of Real-Time Hydrogen Peroxide Dynamics in vivo* [Poster Presentation] Triangle Society for Neuroscience, RTP, NC, USA.
   https://www.trianglesfnchapter.org/_files/ugd/70b47c_ceaa288b748c455d9bc3e098645cfc5f.pdf#page=31

 → *Won the Undergraduate Travel Award for best poster*

4. **Panda, S.**, Riley, S., Wiggins, K., Kathard, R., Alredge, T., & Krause, E. (2018, April). *Developing Solutions for Hand Spasticity* [Poster Presentation] BME Design Symposium, RTP, NC, USA.

5. **Panda, S.**, Wilson, L. R., & Sombers, L. A. (2018, February). *Hydrogen peroxide-specific sensors for In vivo measurements using carbon-fiber microelectrodes* [Poster Presentation] Pittcon, Orlando, FL, USA.

6. Wilson, L. R., **Panda, S.**, & Sombers, L. A. (2017, November). *Hydrogen peroxide-specific sensors for In vivo measurements using carbon-fiber microelectrodes* [Poster Presentation] Society for Neuroscience, Washington, DC, USA.
   https://www.abstractsonline.com/pp8/index.html#!/4376/presentation/19683

7. **Panda, S.**, Wilson, L. R., & Sombers, L. A. (2017, August). *Hydrogen Peroxide Specific Sensors For In Vivo Measurements Using Chronically Implanted Carbon-fiber Microelectrodes* [Poster Presentation] Summer Undergraduate Research Symposium, Raleigh, NC, USA.

8. **Panda, S.**, Wilson, L. R., & Sombers, L. A. (2017, April). *Determining the Sources That Contribute to Extracellular Hydrogen Peroxide Dynamics in the Striatum* [Poster Presentation] Triangle Society for Neuroscience, RTP, NC, USA.
   https://www.trianglesfnchapter.org/_files/ugd/70b47c_42aa665faa94404fb0f52646801378a7.pdf#page=28

9. **Panda, S.**, Wilson, L. R., & Sombers, L. A. (2017, April). *Hydrogen peroxide specific sensors for in vivo measurements using chronically implanted carbon-fiber microelectrodes* [Poster Presentation] Spring Undergraduate Research Symposium, Raleigh, NC, USA.

10. **Panda, S.**, Wilson, L. R., & Sombers, L. A. (2016, December). *Multiple Sources Contribute to Extracellular Hydrogen Peroxide Dynamics in the Striatum* [Poster Presentation] Keck Center for Behavioral Biology Conference, Raleigh, NC, USA.

11. **Panda, S.**, Wilson, L. R., Schmidt, A. C., & Sombers, L. A. (2016, November). *Multiple sources contribute to extracellular $H2O2$ dynamics in the striatum* [Poster Presentation] Society for Neuroscience, San Diego, CA, USA.
    https://www.abstractsonline.com/pp8/index.html#!/4071/presentation/22335

12. **Panda, S.**, Wilson, L. R., & Sombers, L. A. (2016, August). *Multiple Sources Contribute to Extracellular $H2O2$ Dynamics in the Striatum* [Poster Presentation] Summer Undergraduate Research Symposium, Raleigh, NC, USA.

13. **Panda, S.**, Wilson, L. R., & Sombers, L. A. (2016, April). *Determining the Sources That Contribute to Extracellular Hydrogen Peroxide Dynamics in the Striatum* [Poster Presentation] Triangle Society for Neuroscience, RTP, NC, USA.
    https://www.trianglesfnchapter.org/_files/ugd/70b47c_42aa665faa94404fb0f52646801378a7.pdf#page=28

14. **Panda, S.**, Wilson, L. R., & Sombers, L. A. (2016, April). *Determining the Sources That Contribute to Extracellular $H2O2$ Dynamics in the Striatum* [Poster Presentation] Spring Undergraduate Research Symposium, Raleigh, NC, USA.

SKILLS

Python (LangChain, FastAPI, PyTorch, Dash, scikit-learn, pandas, TensorFlow), LLM APIs (OpenAI, Gemini, Vercel AI SDK), SQL (Google BigQuery), Cloud Services (Google, AWS, Azure), React (Next.js), R, Cython, Tailwind CSS, Developer Tools (Git, Docker), Continuous Integration (CircleCI, Travis CI) HTML, MATLAB, Unix Shell Scripts, Familiarity with C/C++, Java

AWARDS & HONORS

Computational Biology Fellowship, *Johns Hopkins University*	**2020**
AWS IMAGINE Grant, *Amazon Web Services* (Supported the mgcpy (now hyppo) package)	**2018**
Magna Cum Laude, *NC State University*	**2018**
University Honors Program, *NC State University*	**2018**
Dean's List, *NC State University*	**2014 – 2018**
Enrichment Grants, *Goodnight Scholars Program, NC State University*	**2014 – 2018**
Goodnight Scholarship, *NC State University*	**2014**
National Merit Corporate Scholarship, *National Merit Scholarship Corporation*	**2014**

TEACHING

NeuroData Design I (EN.580.237/437/697) & II (EN.580.238/438/638) — 2020-21, 2021-22
Johns Hopkins — *Baltimore, MD*
- Formulated projects for students and guided students during weekly presentations.
- Graded students' final projects (which involved code-review of a pull-request).

Computer Methods in Biomedical Engineering (BME 201) — Fall 2017
NC State — *Raleigh, NC*
- Taught multiple lab sections in which students would solve coding problems assigned to them.
- Graded students' code, homework assignments, and tests.

Biomedical Electronics (BME 210) — Spring 2017
NC State — *Raleigh, NC*
- Ran lab sections where students were taught the basics of circuits.
- Graded students' lab quizzes and homeworks

SERVICE

Research

Life Sciences Advisor, *A-Level Capital* — May 2022 – Dec 2024
- Advised students on scientific validity of start-ups and sourced early-stage life sciences companies.

Journal Reviewer, *SoftwareX* — Nov 2022 – Jan 2023

Conference Co-Chair & Reviewer, *SciPy Conference* — 2020, 2021, 2023
- Co-Chair Tracks: Scientific Applications of Biology and Bioinformatics (2020); Biology and Neuroscience (2021), Bioinformatics, Computational Biology, & Neuroscience (2023)
- Reviewer Tracks: Machine Learning and Data Science (2020); Scientific Applications of Machine Learning and Data Science (2021)

Other Service

President, *Ramchandra Panda Scholarship Trust, Balasore, Orissa, India* — Jun 2012 – Present
- Mission: Provide rural students the opportunity to build a better education by providing a monetary investment and by helping preserve traditions for future generations.
- Started as a math competition given to 50 rural students in my home-village in Balasore, Orissa, India, and has grown to 300+ students in both the village and neighboring town specializing in math, traditional Indian dance, art, and music.

Various Leadership Positions, *Goodnight Scholars Program* — Sep 2014 – May 2018
- Committee Chair/Member: Planned yearly scholar events to give back to the community.
- Ambassador: Raised awareness about the scholarship in the local community.
- Mentor: Helped first-year scholars transition to NC State.
- Tutor: Helped younger scholars in various classes.
- Senior Gift Member: Raised money for the state Science Olympiad and served as a judge there.

Volunteer, *Neurosciences Hospital (UNC Healthcare)* — Jul 2017 – Jan 2018
- Assisted nurses and medical staff to improve the safety, comfort, and care of the patients.

Volunteer, *FIMRC (Foundation for International Medical Relief of Children)* — Dec 2017
- Improved health in a village in Kodaikanal, Tamil Nadu, India by taking vital signs of patients and shadowing local physicians.
- Built free chimneys for around 10 residents.