

# Sambriddhi Mainali, Ph.D.

Wildwood MO | 919-525-7041 | mainali.sambriddhi@gmail.com | [LinkedIn](#) | [Publications](#)

---

## Career Summary

---

As an educator, my goal is to create an environment where students develop their skills and knowledge so that they can meet their academic and professional goals through my experience in academia and industry. I am an analytical, innovative, and results-driven Data Scientist with 5+ years of experience in developing and implementing machine learning algorithms to drive business insights and solve complex problems. I have a proven track record of collaborating with cross-functional stakeholders and global teams to deliver impactful solutions.

---

## Professional Experience

---

- **Assistant Teaching Professor (Computer Science Department)**  
University of Missouri-St Louis  
August 2023 - Present
- **Data Scientist (Contractor)**  
Bayer Crop Science  
August 2021 - July 2023
  - Developed and oversaw a variant discovery bioinformatics pipeline within a cloud environment, supporting global pipelines for biological feature prediction requests for regions including North America, India, Pakistan, Thailand, Philippines, Europe, Africa, and more.
  - Collaborated with Lead Data Scientist to
    - successfully process requests and deliver suitable clusters in crops to regions including North America, Africa, and Latin America to help breeders make decisions about seed production
    - successfully provide the prescription for the number of seeds to be planted in different fields in North America
- **Part-time Faculty (Computer Science Department)**  
University of Missouri-St Louis  
January 2023 - August 2023
  - Teaching the course entitled “Python for Data Science and Scientific Computing”
- **Graduate Assistant**  
The University of Memphis  
January 2016 - April 2021

- Collaborated with 10+ scholars on projects involving multiple disciplines, including Biology, Statistics, Computer Science, and Economics.
- Coached and mentored international students on research projects involving machine learning models, species identification in bacteria, virus classification, malware classification, and disease management.
- Conducted research in the field of biological data science in projects to
  - develop prototypes of dimensionality-reduction methods to vectorize DNA sequences, as part of research on DNA vectorization.
  - train and validate deep neural networks for predicting phenotypic features and environmental factors of the habitat of *Arabidopsis thaliana* (Thale grass) and blackflies, as part of research on phenotypic and environmental feature prediction models.
  - develop and validate a prototype for vectorizing DNA sequences of organisms to map them based on their taxonomic association with ~80% accuracy, precision, and recall, as part of research on species delimitation and pathogenicity.
- Assisted in teaching courses including Machine Learning, Fundamentals of Data Science, Introduction to Computer Science, Problem-Solving with Computers, Object Oriented Programming and Data Structures, Foundations of Computing, and grading assignments
- Presented research findings at IWBBIO 2020 virtually (presentations available at <https://www.dropbox.com/sh/k3eh9cdrcw3w2n9/AACVYya0IXPYdHMI7ruO hBLGa?dl=0>)
- Assisted in developing online course content for Fundamentals of Data Science (UM Global)
- Guest Lecture (Invited) Jul 2017
  - Delivered guest lecture on *Towards Reliable Microarray Analysis and Design* at the International Engineering Summer School in Data Science, the National University of Colombia

---

### **Education**

---

- Ph.D.: Computer Science, University of Memphis, 2021
- Master of Science: Computer Science, University of Memphis, 2017
- Bachelor of Science: Computer Science And Information Technology, Kathford International College, 2014

---

### **Technical Skills, Software, & Programs**

---

- Programming: Python, R, SQL, JCL, Java
- Supervised Learning: linear and logistic regressions, decision trees, support vector machines (SVM), random forest, xgboost, neural networks, deep and convolutional networks

- Unsupervised Learning: k-means clustering, hierarchical clustering, k-nearest neighbors, Principal component analysis (PCA), autoencoders
- Data Visualization: Plotly, Matplotlib, Seaborn, Excel, Google Charts
- Cloud Computing Platforms: Amazon SageMaker, Google Cloud: BigQuery, Domino

---

### Awards & Volunteer Experience

---

- **First Prize: Computer Science Research Day**, University of Memphis (2021, 2018)
- **1st Class Ranking (4 Semesters)**, Kathford International College (2017)
- **Volunteer Data Analyst**, Center for Cooperation and Developmental Nepal (2008-2009)

---

### Publications

---

- Garzon, M., & Mainali, S. (2022). *Molecular Computing Approaches. In Dimensionality Reduction in Data Science (pp. 145-167). Springer, Cham.*
- Garzon, M., Mainali, S., & Jana, K. (2022). *Information-Theoretic Approaches. In Dimensionality Reduction in Data Science (pp. 127-144). Springer, Cham.*
- Garzon, M., Mainali, S., Chacon, M. F., & Azizzadeh-Roodpish, S. (2022). *A computational approach to biological pathogenicity. Molecular Genetics and Genomics, 1-14.*
- Garzon, M., and Mainali, S. (2022). *Deep structure of DNA for genomic analysis. Human molecular genetics, 31(4), 576-586.*
- Azzizadeh-Roodpish, S, Garzon, M, Mainali, S. *Classifying Single-Nucleotide Polymorphism in Humans. Molecular Genetics and Genomics, 296(5), 1161-1173.*
- Mainali S, Garzon M, Venugopal D., Et Al. *An Information-theoretic Approach to Dimensionality Reduction in Data Science. Int. Journal of Data Science and Analytics, 12(3), 185-203.*
- Mainali S., Colorado F. A., and Garzon M., "Foretelling the Phenotype of a Genomic Sequence," in *IEEE/ACM Transactions on Computational Biology and Bioinformatics 18:2, pp. 777-783, 1 March-April 2021, doi: 10.1109/TCBB.2020.2985349.*
- Mainali S., Garzon M.H., Colorado F.A. (2020). *Profiling Environmental Conditions from DNA. In: Rojas I., Valenzuela O., Rojas F., Herrera L., Ortuño F. (eds) Bioinformatics and Biomedical Engineering. IWBBIO 2020. Springer Lecture Notes in Bioinformatics 12108, 647–658. doi.org/10.1007/978-3-030- 45385-5.*
- Mainali S., Garzon M.H., Colorado F.A. (2020) *New Genomic Information Systems (GenISs): Species Delimitation and IDentification. In: Rojas I., Valenzuela O., Rojas F., Herrera L., Ortuño F. (eds) Bioinformatics and Biomedical Engineering. IWBBIO 2020. Springer Lecture Notes in Bioinformatics 12108, 163-174. doi.org/10.1007/978-3-030- 45385-5\_15.*
- Garzon. M, Mainali, S. (2017). *Towards a Universal Genomic Positioning System: Phylogenetics and Species Identification. In: Proc. 5th Int. Work-Conference on Bioinformatics and Biomedical Engineering. Springer Lecture Notes in Bioinformatics 10209, 469-479.*
- Garzon, M, Mainali, S. (2017). *Towards Reliable Microarray Analysis and Design. Proc. of the 9th Int. Conference on Bioinformatics and Computational Biology, BiCOB'17, 6 pp*

**Reference available upon request.**