Dr.-Ing. Debdas Paul

Address: Ernst-Toller-Str. 16, 99086 Erfurt, Germany
E-Mail: paul.debdas@gmail.com
Date of Birth: 20.10.1986
Mobile phone: **+49 176 217 444 21**
More about me: Homepage, Linkedin, Scholar, Github, Erdös Number, XING

SUMMARY

I am an **engineer** and a **scientist** by training, so my working philosophy is not only grounded in solution-driven approaches but also in digging deeper into reasoning and generating ideas. I thrive in **interdisciplinary, cross-functional environments**. I began my career as a **computer science engineer**, but my fascination for **complex systems** (biological, financial, communcation etc.) motivated me combine concepts from computer science, control and sysetms theory and applied mathematics and machine learning to understand the intricacies of the system. Primarily, I addressed problems in biological systems so far, but at core I am a computer scientist who thrives on solving complex problems. Notably, biological systems provide an interesting paradigm when compared with other complex adaptive systems. For example, in financial markets, investors continuously adapt, compete, and evolve their strategies in response to changing environments, resembling evolutionary dynamics observed in biological systems rather than the behavior of predictable mechanical systems.

EDUCATION

Stuttgart, Germany — **University of Stuttgart** — **2014 - 2019**

- **Dr.-Ing.** (**PhD in Engineering**) with ***Magna Cum Laude***.
- **Advisor:** Prof. Dr. rer. nat. Nicole E Radde**, Institute of Stochastics and Applications**
- **Technical Skills**: MATLAB, Numerical simulation, Systems Theory, ODE based modeling, Stochastic simulation, ARIMA models

Finland and Stockholm, Sweden — **Aalto University and KTH** — **2012 - 2014**

- **MS** and **MSc. (Tech.)** in **Computational Systems Biology** with a ***Distinction***
- **Technical Skills**: MATLAB, Python, R
- **Selected subjects**: Machine learning, Applied and numerical mathematics

Kolkata, India — **Jadavpur University** — **2009 - 2011**

- **Master of Engineering** in **Computer Science & Engineering**
- **Technical skills**: Python
- **Selected subjects**: Advanced algorithm design, Machine learning, Information theory, NLP, Pattern recognition

Kolkata, India — **West Bengal University of Technology** — **2005 - 2009**

- **Bachelor of Technology** in **Computer Science & Engineering**
- **Technical skills**: C, Java
- **Selected subjects**: Algorithm design, Theoretical Computer Science, Operating systems, Engineering mathematics, Control systems, Engineering physics, Database management systems.

EMPLOYMENTS

Postdoc. Data scientist/Project leader/Tech lead — **Leibniz Institute on Aging, Jena, Germany** — **April 2023 - present**

- **Problem definition**: **Predictive models** often fail to generalize well in out-of-distribution (OOD) settings. On top of that, predictions frequently suffer from unwanted bias, resulting in unfair and overly optimistic solutions. Irrespective of domain, this has large implications when a solution is affecting millions of lives. The

question is: *can we develop a principled, unified approach that tackles the competing objectives of **OOD generalization**, **bias mitigation** and **fairness**, and **explainability** (when using deep neural networks)?*

- **Solution approach**: **Representation learning** via **Adversarial training** learns a latent representation that can be made bias free and predictive using a min-max objective. While implementing using deep neural network, a l1- based layer at the beginning can dynamically extract important features (explainability via feature attribution).
- **My contribution:** Implementation of the model architecture using deep neural network, fine-tunining, theoretical analysis of the scope and limitation of such approach compare to classical machine learning models (regression based, tree-based), demonstrating practical applicability using moderately high-dimensional(5-6k features), noisy molecular data [Pre-print on request]
- **Tech stack used:** Python, Tensor-Flow, Scikit-learn
- **From idea to product:** Won a grant worth ~100K EUR from the German Federal Ministry for Research, Technology and Aeronautics, leading the technical side to bring this into a product with a goal to disocver novel pharmacological interventions for aging, reducing time and cost **[media coverage].**
- **Dissemination**: This work is **selected for a spotlight presentation and poster** at the **EurIPS Causality for Impact workshop, 2025** in Copenhagen.

Postdoctoral scientist (machine learning) — **University Hospital, Tübingen, Germany** — **June 2020 - August, 2022**

- **Problem definition:** Consider tabular data that are order-invariant, meaning that shuffling the rows does not affect the underlying structure, unlike image data where spatial ordering is essential. The objective is to identify groups of data points, represented by n-dimensional feature vectors, that can statistically and significantly distinguish between two groups in a supervised setting. *A key question is whether an appropriate data-driven approach can be developed when the number of such data points is very low.*
- **Solution approach:** A weakly supervised representation learning framework implemented using CNNs learns to identify low-frequency discriminative points, which clustering-based approaches typically fail to detect.
- **My contribution**: I contributed as a computational expert in a multi-institutional collaborative project including expert and non-expert, adapting and applying an in-house model (based on the solution above) to extract meaningful patterns tailored to the problem domain. I worked in the domain of auto-immune diseases.
- **Dissemination:** The findings were disseminated in *Frontiers in Immunology*.

Postdoctoral scientist — **Max Planck Inst. for Biophysical Chemistry, Germany** — **Feb 2019 - May, 2020**

- **Problem definition**: Our immune system, which protects us from all sorts of infection, is an extremely robust system that can be imgained as a pattern recognition system. The system identfies patterns displayed by cells and then the system tags those cells with healthy or un-healthy status. These patterns are generated by elements called *peptides*, generated by an unit called Proteasome. During Cancer, this pattern recognition system fails. Accurate quantification of peptides helps to devise targated therapy for cancer. The question is, *can we accuratly absolutely qunatify the peptides without further time and cost-inefficient experimentations?*
- **Solution approach**: The absolute qunatification requires estimation of a multplicative factor that converts signals to concentration (quantification) directly without further experiment.
- **My contribution:** I contributed to the initial design and implementation (in R language) of an algorithm based on Bayesian Inference using MCMC sampling scheme, that estimate this multiplicatibe factor.
- **Dissemination:** The findings were disseminated in *Frontiers in Immunology*.

Doctoral researcher **University of Stuttgart, Germany** **Nov 2014 - April, 2019**

- **Problem**: Biological systems reliably transmit and process signals in the presence of significant noise and non-stationary inputs. But how it achieves this? How the underlying design principle contribute to such robustness?
- **Applied** concepts from **systems theory, non-linear dynamcis, stochastics** to uncover key mechanisms that sustain signal propagation and resilience in biological systems.
- **Leveraged systems theory** and performed **stochastic simulations** to analyze **signal propagation dynamics**.
- Focused on the role of retroactive effects (bi-directional signal propagation) in maintaining the resilience of phosphorylation cascades.
- **Discovered** the critical role of retroactive effects in sustaining resilience within a phosphorylation cascade, contributing to a deeper understanding of biological signal propagation.
- **Disseminated** the findings in a book chapter*, a conference proceeding*,#, and in the *Journal of Theoretical Biology**,#. *=first author, # = corresponding author

PROJECTS AS A VISTING RESEARCHER AND AS PART OF INDEPENDENT COLLABORATION

Boston, MA, USA **Dept. of Systems Biology, Harvard Medical School** **07/2017 - 10/2017**

- **Developed** a rule-based model using the Kappa language framework to study gene regulation.
- **Created** a succinct and adaptable graphical representation of molecular interactions.
- **Provided** valuable insights into the dynamics of gene regulation and molecular processes.
- **Advisor: Prof. Jeremy Gunawardena**

Uppsala, Sweden **Division of Scientific computing, Uppsala University** **01/2014 - 06/2014**

- **Conducted** numerical analysis to evaluate an optimized pre-conditioning scheme.
- **Demonstrated** that the scheme significantly **reduced computation steps** for exact Chemical Master Equation (CME) calculations by **approximately 80%** while maintaining comparable convergence characteristics despite the reduction in computation steps.
- Results were disseminated in the form of masters thesis in theoretical biological physics
- **Advisor: Prof. Stefan Engblom**

Koper, Slovenia **Dept. of Mathematics, University of Primorska** **02/2011 - 03/2011**

- **Implemented** eigenvector-based algorithms to identify large bipartite sub-graphs.
- **Utilized** sign patterns of eigenvectors to enhance the accuracy of the algorithm.
- **Achieved** results that are consistent with Erdős' bound for the graph.
- **Published** the findings in *Discrete Applied Mathematics*[*,#]. [TALK], *=first author, # = corresponding author
- **Collaborator: Prof. Dragan Stevanovic**

KEY PROGRAMMING AND TECHNICAL SKILLS

- **Languages**: **Python** [**for machine learning, frequently used**], R [**for data visualization**], MATLAB [**used as doctoral student for ODE based modeling and stochastic simulation**]
- **Machine learning/Deep learning libraries**: Scikit-learn, TensorFlow [**related to current position at the Leibniz Institute and the postdoc. position at the University Hospital Tübingen**]
- **Machine learning models**: Regression based **(K-NN, Linear/Logistic regressor, ElasticNet),** Tree-based models (**XGBoost, LightGBM**) - **for benchmarking purpose against neural network-based models.**
- **Data manipulation/visualization:** Pandas, NumPy, Matplotlib, Seaborn [**related to current position at the Leibniz Institute**]
- **Version control/container technologies**: Git [**frequently use for my projects**], Docker [**seldom use it**]

AWARDS

- Go-Bio initial from the **Federal Ministry for Research, Technology and Aeronautics**, Germany ~100K EUR for 2024-25
- **European Union's Erasemus Mundus Fellowship**, equivalent to Fullbright, totaling EUR 48K for 2012-14.
- **Bilateral Mobility Grant** from the Government of the Republic of Slovenia in 2011.

INVITED TALKS

- Comparison between combinatorial and spectral approaches in identifying the largest bipartite subgraphs of a graph at the Workshop on Graph Spectra, Combinatorics, and Optimization – WGSCO 2018, Aveiro, Portugal[SLIDES].
- Stochastic sequestration dynamics can act as an intrinsic noise filter in signaling network motifs at INRIA Saclay - Île-de-France research centre, 2018[SLIDES]

TEACHING AND ORGANIZATIONAL DUTIES

- Course assistant: **Statical learning methods and stochastic control systems**, graduate course, University of Stuttgart, Institute for Systems Theory and Automatic Control, 2017
- Course assistant: **Systems theory in Systems biology**, graduate course, University of Stuttgart, Institute for Systems Theory and Automatic Control, 2016
- Introduction to MATLAB: Course organized for bachelor students in the Engineering Faculty at the University of Stuttgart. (Winter semester 2018/19)

PRIVATE CONTINUOUS LEARNING

Whenever I have time, I try to stay updated on technological advancements relevant to the software industry. (clickable):

- Oracle Cloud Infrastructure 2025 Certified AI Foundations Associate (Oct, 2025) - BADGE
- End-to-end machine learning operations (MLOps) with Azure Machine Learning [July, 2023]
- Microsoft Azure Fundamentals: Describe cloud concepts [July, 2023]
- Deploying Scalable Machine Learning for Data Science [Jan, 2022]
- Succeeding in DevOps [Jan, 2022]

LANGUAGES I SPEAK

- English (Fluent) - German (B1, Goethe) - Bengali (Native), -Hindi (Fluent)

ACADEMIC REFERENCES

1. Name: **Prof. Dr. rer. nat. Nicole Radde**

Affiliation: Institute for Stochastics and Applications, University of Stuttgart, Germany
Relation: Doctoral advisor

Email: nicole.radde@simtech.uni-stuttgart.de

2. Name: **Prof. Alessandro Cellerino**

Affiliation: Associate Professor of Physiology Bio@SNSScuola Normale Superiore, Pisa, Italy & Leibniz Institute on Aging -Fritz Lipmann Institute Jena, Germany
Relation: Post-doctoral advisor

Email: alessandro.cellerino@sns.it

3. Name: **Prof. Stefan Engblom**

Affiliation: Divison of Scientific Computing, Uppsala University, Sweden
Relation: Masters thesis advisor
Email: stefan.engblom@it.uu.se

4. Name: **Prof. Dragan Stevanovic**

Affiliation: Abdullah Al Salem University, Kuwait
Relation: collaborator
Email: dragan.stevanovic@aasu.edu.kw