

Rijswijk, The Netherlands

📞 (+31) 613485171

✉ sayakmukherjee2010@gmail.com

🌐 sayakmukherjee.github.io

in sayakmukherjee1996

🔗 SayakMukherjee

# Sayak Mukherjee

## Experience

09/2023–Present **Researcher, TU Delft, Delft, NL**

- Worked on generative models with a focus on diffusion models.
- Focused on developing methods for training-free alignment of diffusion models.
- Proposed a sampling-based alignment approach based on importance sampling.
- Collaborated with researchers from Shell AI Research and Google Deepmind.

09/2018–08/2021 **Application Developer, Oracle, Bangalore, IN**

- Collaborated with financial stakeholders to gather requirements.
- Customized Oracle FLEXCUBE core banking platform for Banco de Occidente (Colombia) and Rabobank (Brazil), optimizing for regional banking needs.
- Developed 40+ functionalities tailored to the Latin American banking sector.
- Utilized Java, Spring, Hibernate, SOAP API, and JavaScript.
- Deployed solutions on Oracle Database and Oracle Weblogic Server.

07/2017–08/2017 **Data Science Intern, Indian Institute of Management, Lucknow, Remote**

- Covered basics of R programming, data analytics, statistics, and econometrics.
- Used acquired knowledge to solve business cases by Harvard Business School.
- Completed a capstone project on hotel room pricing in Indian market.

## Education

09/2021–08/2023 **MSc Computer Science (Cum Laude), TU Delft, Delft, NL**

*Master Thesis:* Model Agnostic Peer-to-peer Learning

*Key Areas:* Deep Learning, Computer Vision, Distributed Machine Learning

08/2014–07/2018 **BTech Information Technology, WBUT, Kolkata, IN**

*Bachelor Thesis:* Human Action Recognition Using Convolution Neural Networks

*Key Areas:* Data Structures and Algorithms, Soft Computing, Theory of Automata

## Skills

Concepts	Machine Learning, Deep Learning, Computer Vision, Generative AI
ML Libraries	Matplotlib, NumPy, Pandas, SKLearn, TensorFlow, Pytorch, HuggingFace
Languages	Python, Java, R, C, C++
Clouds Services	OCP, GCP, AWS, Azure
MLOps & DevOps	MLFlow, Docker, Kubernetes, Weights & Biases, Git, Streamlit, Gradio
Scripts	HTML, CSS, JavaScript, ReactJS
Soft-skills	Communication, Presentation, Collaboration, Problem-Solving
Software Dev	SDLC, Agile, Scrum, Jira, Confluence, Trello, ClickUp

---

## Publications

- [1] MAPL: Model Agnostic Peer-to-peer Learning, ArXiv '24. [\[Paper\]](#) [\[Code\]](#)  
**Sayak Mukherjee**, Andrea Simonetto, Hadi Jamali-Rad
- [2] Controlled Denoising For Diffusion Models, Under Review.  
**Sayak Mukherjee**, Anuj Singh, Ahmad Beirami, Hadi Jamali-Rad

---

## Academic Activities

Teaching Asst. Human-Computer Interaction ([CSE3500](#)), Web Science & Engineering ([IN4252](#)), Fundamentals of AI Program ([IFEEMCS520100](#)), Deep Learning ([CS4240](#)), AI Skills for Engineers ([EdX](#))

---

## Other Courses

- 07/2024 **Oxford Machine Learning Summer School, AI for Global Goals**
  - Completed a 22-hour program on [Representation Learning and Generative AI](#).
  - Studied neural and behavioural comparisons between humans and machines.
  - Learned methods for uncertainty quantification in AI systems.
  - Investigated representation learning and generative AI for computer vision tasks.
- 06/2024 **Generative Modelling Summer School, TU Eindhoven, TU Denmark, Inria**
  - Completed 30.5 hours of training, including lectures, exercises, and a poster session.
  - Gained specialized knowledge in generative modelling.
  - Engaged with interdisciplinary applications such as bioinformatics, computational physics, and social science.
  - Full program is available at [gemss.ai/2024](#).
- 08/2020 **Machine Learning, Stanford Online via Coursera**
  - Learnt about Linear Regression, Logistic Regression and Regularization.
  - Covered Machine Learning System Design, Supervised and Unsupervised Learning.
- 07/2020 **Microsoft Azure Machine Learning Foundation, Microsoft via Udacity**
  - Implemented Model Training on Azure Machine Learning Studio.
  - Covered Feature Engineering, Feature Selection and Data drift.
  - Learnt about Supervised, Unsupervised and Reinforcement Learning.
  - Studied concepts such as Model Transparency, Explainability, and Fairness.
- 05/2020 **Data Engineering with Google Cloud Specialization, Google via Coursera**
  - Covered the fundamentals of Big Data and AI on Google Cloud Platform.
  - Learnt about modernization of Data Lakes and Data Warehouse.
  - Used Cloud Dataproc, Cloud Dataflow, Cloud Data Fusion and Cloud Composer to build and manage data pipelines.
  - Built streaming analytics systems using Cloud Pub Sub, Big Query and Big Table.
  - Learnt about Machine Learning using Cloud AutoML, Big Query ML and Kubeflow.

---

## Languages

**English** (Professional Working Proficiency; CEFR Level: C1, IELTS dt. 06/2019), **Hindi** (Professional Working Proficiency), **Bengali** (Native Proficiency)

---

## Hobbies

**Singing** (Certificate in Rabindra Sangeet), **Painting** (Senior Diploma), **Playing piano and guitar**