



Sarbashis Das
Email: dsarbashis@gmail.com

in: www.linkedin.com/in/sarbashis/

Q: <https://sarbashis.github.io>
☎ (+46) 765-770649

Professional Skills

Project Management

- JIRA
- Sprint Planning/Review
- Risk Management

Team Leadership

- Cross-functional team
- Support & Mentoring

Machine Learning

- Deep Learning
- Natural Language Processing (NLP)
- Principle Component Analysis (PCA)

Machine Learning Framework

- Keras
- Tesorflow
- scikit-learn
- Spacy

Deployment & Monitoring

- Docker
- Kubernetes & Grafana
- Elasticsearch & Kibana

Data Modelling

- Knowledge Graph
- Ontology

Data Mining

- Information Extraction
- Airflow DAG

Programming

- Python
- R

Address

Räcksta Gårdsväg 28
16273, Vällingby
Stockholm, Sweden

Sarbashis Das, PhD.

Technical Project Manager & Senior Data Scientist

About Me Experienced Senior Data Scientist and Project Manager with a PhD focused on product development using various Machine Learning methods. Skilled in Machine Learning and Artificial Intelligence with a proven track record of success leading diverse projects of various types. Strong analytical and problem-solving skills, combined with technical knowledge, have enabled me to develop innovative solutions that consistently exceed expectations.

Work Experience

2020 – Present: *Technical Project Manager & Senior Data Scientist*

Global AI Accelerator (GAIA) Ericsson, Sweden

- **Project: Development of AI based Virtual Assistant (Ericsson Production Information Assistant)**
 - Tasks:
 - * Managed a team of 15 individuals including Data Scientists, Data Engineers, Software Developers, and Domain Experts to deliver a successful project.
 - * Awarded **Ericson Impact Award** for the strong performance and business impact.
 - * Implemented **Agile Methodologies** for planning and managing the development of the Assistant.
 - * Conducted Sprint Planning and Sprint Review, and effectively communicated project status to stakeholders.
 - * Contributed to the development of the **product roadmap**, managed resources, and oversaw budgeting.
 - * Responsible for the overall the Architecture of the Assistant
 - * Designed and implemented a cloud-based infrastructure using **AZURE Kubernetes** and monitored performance using **Grafana**.
 - * Set up an automated data processing pipeline using **Airflow** and collaborated with the infra-team for automated deployment in Kubernetes using **GitOps ArgoCD**.
 - * Proposed and implemented solutions for automated data mining and knowledge graph generation.
 - * Proposed and drove the implementation of **Semantic Search** into the system.
 - * Provided support to team members and mentored junior members and interns.

2019 – 2020: *Data Scientist*

Global AI Accelerator (GAIA) Ericsson, Sweden

- **Project: Development of AI based Virtual Assistant (Ericsson Production Information Assistant)**
 - Tasks: **Information retrieval** and **knowledge graph** building from Product Documents
 - * **Dataset:** Text Documents, structured, semi-structured and unstructured
 - * **Tools:** **Spacy, Stanford CoreNLP, Python, Flask**
 - * Developed pipeline for data mininnng documents and knowledge graph generation using OpenIE.
 - * Designed **SPARQL** queries to fetch information from the knowledge graph
 - Developed micro-service using flask and Cellery to schedule job triggering for regular testing.
 - Supervised Master thesis student.

2016 – 2019 *Senior Researcher*

Scilifelab, Uppsala University, Sweden

- **Understanding gene regulation in different types of heart failures**
 - Lead research project to identify genes & regulatory mechanisms in different types of Heart Failures and published in Nature Scientific Report.
 - Dataset: High-dimension gene expression data (RNASeq) from different Heart tissues from heart failure patients.
 - Analysis: Dimension reduction using PCA, unsupervised clustering, identification of dysregulated genes.
 - **Tools:** **R, Python, Pandas, seaborn, git**

2014 – 2016, *Researcher*

Uppsala University, Sweden

- **Deep Learning based Protein Function Prediction** Scilifelab, Uppsala University, Sweden
 - **Role:** **Senior Research Scientist**
 - Dataset: Protein Sequences of total 35 millions from public databases
 - Prediction of protein functional class using Deep Neural Network.
 - **Tools:** **Python, Keras, Tensorflow, Jupyter, git**
- **Understanding of Mycobacterial genes and gene regulation using Genomics and Transcriptomic data** Uppsala University, Sweden
 - Dataset: Bacterial Genome & Transcriptome Sequences
 - Analysis: Development of pipeline for analysis, annotation and genome comparison
 - **Tools:** **Python, R, git**



Sarbashis Das
Email: dsarbashis@gmail.com

in: www.linkedin.com/in/sarbashis/

Q:
<https://sarbashis.github.io>
☎ (+46) 765-770649

Professional Skills

Project Management

- JIRA
- Sprint Planning/Review
- Risk Management

Team Leadership

- Cross-functional team
- Support & Mentoring

Machine Learning

- Deep Learning
- Natural Language Processing (NLP)
- Principle Component Analysis (PCA)

Machine Learning Framework

- Keras
- Tesorflow
- scikit-learn
- Spacy

Deployment & Monitoring

- Docker
- Kubernetes & Grafana
- Elasticsearch & Kibana

Data Modelling

- Knowledge Graph
- Ontology

Data Mining

- Information Extraction
- Airflow DAG

Programming

- Python
- R

Address

Råcksta Gårdsväg 28
16273, Vällingby
Stockholm, Sweden

2013 – 2014, *Postdoctoral Researcher*

Institute of Molecular Biology, Germany

- **Understanding of chromatic accessibility during development**
 - Dataset: FAIRseq & Transcriptome Sequences
 - Analysis: identification of genomic regions with open chromatin and gene regulation
 - Tools: FASTQC, R, Plotly

Certifications

I. Specialization in Deep Learning

It consists of five courses from Coursera

II. Specialization in Machine Learning Engineering for Production (MLOps).

It consists of four courses from Coursera

II. Architecting with Google Kubernetes Engine: Workloads

DeepLearning AI, Coursera

III. Architecting with Google Kubernetes Engine: Foundations

DeepLearning AI, Coursera

IV, Machine Learning A-Z TM Hands-On Python & R In Data Science

From Udemy

V. Machine Learning

Stanford University, Coursera

VI. Introduction to Data Science in Python

University of Michigan, Coursera

Awards

2021, Ericsson Impact Award 2021

The award was given to the people who contribution has a great impact to Ericsson business value

2017, Research Grant awarded as a co-applicant

Title: Dynamics of gene regulatory networks in response to different stresses in Mycobacteria.
Funding Agency: FORMAS.

2016, Research Grant awarded as a co-applicant

Title: Non-coding and regulatory RNAs in Mycobacteria. Funding Agency: Swedish Research Council (VR)

2012, Awarded Wennergren Postdoc Fellowship

Funding Agency: Wennergren Foundations. Grant offered, but gracefully declined due to accepting other Postdoc position)

2010, Awarded Senior Research Fellowship

Funding Agency: Council of Scientific and Industrial Research (CSIR), Government of India.)

2006, Awarded Scholarship for Master of Technology

Funding Agency: Department of Biotechnology (DBT), Government of India.)

Education

2009 - 2013, Ph.D. in the field of Data Science (Biological Data)

Jawaharlal Nehru University (JNU), India.

2006 - 2008, M.Tech in the field of Data Science (Biological Data)

Jawaharlal Nehru University (JNU), India.

2003 - 2005, Burdwan University (BU), India.

M.Sc in Zoology

Publications

I have a total 14 publications in reputed peer-reviewed journals with many first authors.
For more information check Google Scholar page.

Communication Skills

2022, Invited talk, Uppsala University

AI/ML and it's application in Biology to Master Program students

2022, Presented & demoed, Ericsson Developer Conference (Stockholm)

Presented Ericsson Product Information Assistant and demoed different functionalities

2022, Data and AI submit 2022- Databricks

Interests

Professional

AI, Machine Learning, MLOps, NLP, Software development

Personal

Photography, Cooking