europass

Suparno Datta

Date of birth: 12/12/1989 | **Nationality:** German | **Gender:** Male | **Phone number:** (+49) 17632433883 (Mobile) | **Email address:** suparno.datta@gmail.com | **Website:** https://scholar.google.de/citations?hl=en&user=v6xGwm8AAAAJ | **LinkedIn:** linkedin.com/in/suparno-datta-70573a2a |

Address: Pasewalker Straße 72, 13127, Berlin, Germany (Home)

WORK EXPERIENCE

01/03/2024 – CURRENT Berlin, Germany

TECH LEAD DATA SCIENCE FRESENIUS MEDICAL CARE

1. Leading a cross-functional team of data scientists, engineers, and analysts in the development and implementation of advanced optimization based solution which aims to keep a patient's hemoglobin within a specific range by recommending correct drug doses. I also serve as the lead architect for the project, overseeing the complete design of the architecture on the Azure platform and ensuring readiness for FDA submission.
2. Designed and Implemented predictive modeling projects to anticipate patient needs and outcomes, enhancing proactive care strategies and reducing hospital readmissions.
3. Provided mentorship and technical guidance to junior team members, fostering a culture of continuous learning and innovation within the data science team.
4. Communicated complex analytical insights and model results to non-technical stakeholders through clear and actionable reports and presentations, facilitating data-driven decision making at all levels of the organization.

01/06/2021 – 29/02/2024 Berlin, Germany

SENIOR DATA AND APPLIED SCIENTIST MICROSOFT

1. Engineered and deployed a variety of LLM-based applications tailored to specific business and research needs, enhancing functionality and user experience.
2. Utilized Retrieval-Augmented Generation (RAG) techniques to create robust, context-aware systems that integrate external knowledge bases, improving accuracy and relevance in generated responses.
3. Designed and implemented solutions for querying structured data using LLMs, enabling natural language interactions with databases, streamlining data retrieval processes, and making data insights more accessible to non-technical users.
4. Designing, developing and deploying state of the art Machine learning (ML) solutions for clients in various sectors (Health, Manufacturing, and Travel) in Azure.
5. Leading small teams of data scientists and engineers to build end-to-end ML systems and products. Won RISE award, 2023 for outstanding team accomplishment and cross team collaboration.
6. Conceptualizing and developing internal IPs in the area of LLMs, Model explainability and ML Visualization.
7. Tools and Technologies used: Azure ML, Python (various ML and DL libraries such as Pytorch, Keras, Scikit, LangChain, LlamaIndex), Docker (containerization), Azure DevOps (CI/CD).

05/2018 – 31/05/2021 Potsdam, Germany

RESEARCH ASSISTANT, PHD CANDIDATE HASSO PLATTNER INSTITUTE (UNIVERSITY OF POTSDAM)

1. Combining Electronic Health Records, wearable sensor data and Genomic data and applying machine learning methods to better predict individual patient outcomes, patient risk stratification etc.
2. Continuous blood pressure prediction from PPG signals collected from wearable sensors using Convolutional Neural Networks (CNNs).
3. Prediction of health outcomes and onset of diseases from longitudinal sparse electronic health records (EHRs) using approaches like Long Short-Term Memory networks (LSTMs), CNNs etc. and comparing them to other ML algorithms such as XGboost and LightGBM.
4. Comparing different databases (eg: SAP HANA) for optimizing EHR query speeds.
5. Supervising and managing individual master thesis students and master projects performed in groups.
6. Giving lectures to graduate students on various Machine and Deep learning topics.

12/2017 – 05/2018 Munich, Germany
DATA SCIENTIST SCOUT24

1. Develop a model for price prediction of used cars
2. Develop a model for automatic classification of emails for customer care (using word2vec)
3. Programming Languages and tools used - Python, R, Keras, LightGBM

05/2016 – 11/2017 Frankfurt, Germany
DATA SCIENTIST ZEROG GMBH (LUFTHANSA SYSTEMS)

1. Designing and implementing machine learning models (Linear Models, Random Forest, Gradient Boosting Trees etc.) to predict customer affinities for different flight related ancillaries and destinations.
2. Collaborative Filtering for destination recommendation.
3. Designing and implementing various statistical tests to measure campaign success.
4. Working on a variety of descriptive analytics tasks to find valuable insights from the data (Eg: customer segmentation, trend analysis etc.)
5. Present the results to business stakeholders and other data scientists using tableau dashboards.
6. Programming languages and tools used -R, H2O, Python, Teradata SQL, Spark (MLlib), Tableau, Hive.

06/2015 – 04/2016 Göppingen, Germany
DATA SCIENTIST TEAMVIEWER GMBH

1. Working on a wide variety of (predictive) analytics task such as time series modeling for revenue prediction, customer cohort analysis, predicting churn rates, repeat purchase probability prediction etc.
2. Design and develop scalable methods for processing huge amounts of data within a data-driven company to support important business decisions.
3. Programming languages used - R, Java
4. Tools & Technologies used -SQL, Tableau, Hadoop, Impala, Apache Crunch

EDUCATION AND TRAINING

05/2018 – 09/2024 Potsdam, Germany
DR. RER. NAT. IN COMPUTER SCIENCE Hasso Plattner Institute, University of Potsdam

Thesis title: Machine Learning for Early Detection, Management and Prognosis of Hypertension.
Thesis defended successfully on 16.09.2024.

Grade: Magna Cum Laude

08/2012 – 05/2015 Aachen, Germany
MSC. IN MEDIA INFORMATICS Rheinisch-Westfälische Technische Hochschule (RWTH)Aachen, Germany

Principal Courses - Machine Learning, Data Mining, Pattern Recognition, Designing of Interactive systems, Theory of Distributed Systems

Final grade 1.5

08/2007 – 05/2011 Kolkata, India
B TECH IN COMPUTER SCIENCE & ENGINEERING Heritage Institute Of Technology(West Bengal University of Technology)

Principal Courses - Data Structure, Algorithms, DBMS, Engineering mathematics, Networking, Socket Programming, Software engineering.

Final grade 8.2 (out of 10)

LANGUAGE SKILLS

Mother tongue(s): **BENGALI**

Other language(s):

	UNDERSTANDING		SPEAKING		WRITING
	Listening	Reading	Spoken production	Spoken interaction	
ENGLISH	C2	C2	C2	C2	C2
GERMAN	B2	B2	B2	B2	B2

Levels: A1 and A2: Basic user; B1 and B2: Independent user; C1 and C2: Proficient user

PUBLICATIONS

Selected Publications

1. Datta S, Morassi Sasso A, Kiwit N, Bose S, Nadkarni G, Miotto R, Böttinger EP. Predicting hypertension onset from longitudinal electronic health records with deep learning. JAMIA open. 2022 Dec;5(4):ooac097.
2. Hackl M, Datta S, Bottinger E. Unsupervised Learning to Subphenotype Heart Failure Patients from Electronic Health Records. In Artificial Intelligence in Medicine: 19th International Conference on Artificial Intelligence in Medicine, AIME 2021, Virtual Event, June 15-18, 2021, Proceedings (Vol. 12721, p. 219). Springer Nature.
3. Datta S, Sachs JP, Cruz HF, Martensen T, Bode P, Sasso AM, Glicksberg BS, Böttinger E. FIBER: enabling flexible retrieval of electronic health records data for clinical predictive modeling. JAMIA Open. 2021 Jul;4(3)
4. Sasso AM, Datta S, Jeitler M, Steckhan N, Kessler CS, Michalsen A, Arnrich B, Boettinger E. HYPE: Predicting Blood Pressure from Photoplethysmograms in a Hypertensive Population. In International Conference on Artificial Intelligence in Medicine 2020 Aug 25 (pp. 325-335). Springer, Cham.
5. Datta S, Schraplau A, Da Cruz HF, Sachs JP, Mayer F, Böttinger E. A Machine Learning Approach for Non-Invasive Diagnosis of Metabolic Syndrome. In 2019 IEEE 19th International Conference on Bioinformatics and Bioengineering (BIBE) 2019 Oct 28 (pp. 933-940). IEEE.
6. Datta S, Dutta A, Chaudhuri SG, Mukhopadhyaya K. Circle formation by asynchronous transparent fat robots. In International Conference on Distributed Computing and Internet Technology 2013 Feb 5 (pp. 195-207). Springer, Berlin, Heidelberg.
7. Dutta A, Chaudhuri SG, Datta S, Mukhopadhyaya K. Circle formation by asynchronous fat robots with limited visibility. In International Conference on Distributed Computing and Internet Technology 2012 Feb 2 (pp. 83-93). Springer, Berlin, Heidelberg.

HONOURS AND AWARDS

Honours and awards

1. Best Student Paper Award at: Artificial Inteligence in Medicine (AIME), 2020
2. Final year project i.e Circle Formation by Asynchronous Transparent Fat Robots was selected as MOST INNOVATIVE STUDENT PROJECT (Bachelors level), 2011 by The Indian National Academy of Engineering (INAE).