Rajat Kumar

ML Engineer

+91-9639146920

rajat-kumar.netlify.app/

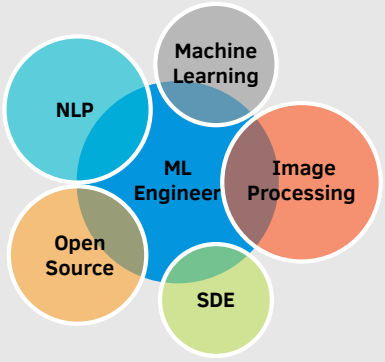
rajat.tech.002@gmail.com

/in/rajat-kumar-543b8a12b

rajat-tech-002

Technical Skills

Overview



Programming

Python • C • C++

Pytorch • Keras

Docker • Kubernetes • AWS

SageMaker • Colab • Java • SQL

Education

M. Tech, ICT (CGPA: 8.18)
Specialization: Machine Learning
Dhirubhai Ambani Institute (DA-IICT)
2018 - 2020 | Gandhinagar, India

B. Tech CSE (CGPA: 7.21)
Gurukul Kangri University
2013 - 2017 | Haridwar, India

Intermediate/+2 (Percentage: 94.8%)
High School (Percentage: 95%)
Lord Mahavira Academy
2010 - 2013 | Saharanpur, India

Professional Summary

With over 4+ years of experience in Research and Development, 250+ google scholar citations, and a master's degree in Machine Intelligence, I currently serve as a data scientist at Philips Innovation Strategy, where my work is driven by a commitment to enhancing patient outcomes. Collaborating closely with clinical scientists and data architects, I analyze complex numerical, textual, and image datasets to develop machine learning models that prioritize patient impact.

Previously, I worked as a researcher at TCS Innovation Lab, focusing on developing NLP-based models for intent detection and discovery using large language models (LLMs). Additionally, I gained hands-on experience during my internship at IIT Gandhinagar, where I worked on the development of energy disaggregation algorithms, specifically utilizing the Non-Intrusive Load Monitoring Toolkit (NILMTK) for smart buildings.

Experience

Oct 2022 - Present — **Data Scientist II - Philips**

- My work centers on data science projects that analyze numerical, text, and image-based patient data. Currently, I am focused on motion correction in MRI Spine images to improve imaging accuracy, leveraging advanced data science techniques for impactful, patient-centered insights. I have also developed data-driven dashboards—such as the Project Ranking Tool and IGM Cost Analysis.

Sep 2020 - Sep 2022 — **Researcher - TCS Innovation Labs**

- Worked in the NLP subgroup of the Deep Learning and AI group.
- Have published and patented my work on the Intent Detection and Discovery Problem, which was presented as a main track at the highly-regarded NAACL (A rated) Conference.
- Tools: Google-Colab, Python, Pytorch, Keras, Jupyter Notebook, GitHub, Docker

May 2019 - Jul 2019 — **Summer Research Intern - IIT Gandhinagar**

- Guide: Dr. Nipun Batra (Assistant Professor at IIT-GN)
- Worked on the open-source toolkit NILMTK (Non-Intrusive Load Monitoring Toolkit) on GitHub, which aims to improve the interface for energy disaggregation problems.
- Tools: Google-Colab, Python, Jupyter Notebook, GitHub
- Published a paper in ACM Buildsys 2019.
- Upgraded the NILMTK library in GitHub.

May 2016 - Jul 2016 — **Summer Intern at Raman Classes, Roorkee**

- Guide: Dr. Ankush Mittal (PhD. at NUS Singapore)
- Worked on Research Based Projects.
- Understood basic ML and Statistics.

Projects

- **Motion Correction in MRI spine images**
  - Simulated artifact induced dataset (like Voluntary and Involuntary patient motion) to create training and test datasets , incorporating realistic artifacts to effectively train deep learning models for enhanced performance in motion correction. Working in collaboration with Clinical scientist and DS architect.
- **Innovation Efficiency: Innovation Outcomes and Analytics**
  - Worked on creating business-specific KPI views and utilizing industry benchmarks and AI to predict/forecast important parameters which help businesses drive the efficiency at the enterprise and business level.

- **Central PMS Data Lake Platform (Philips) (AWS)**
  - Worked together with Research Scientist and Data Engineer to create a Data Lake using AWS for storing PMS (Post Market Surveillance) Data from various sources within Philips.
  - Deployed an outlier detection algorithm using AWS Sagemaker and Docker.
- **Modeling Performance and Power on Disparate Platforms (Open Source)**
  GitHub Link
  - Focused on prediction of performance and power given the CPU architecture and memory features using transfer learning.
  - Worked under the supervision of Prof. Amit Mankodi and co-supervised by Dr. Amit Bhatt(Associate Professor at DA-IICT).
- **NILMTK Contrib Library (Open Source)**
  GitHub Link
  - Created a high level API in nilmtk-contrib (GitHub) which runs Disaggregation algorithms as an addition to NILMTK toolkit.
  - Focused on Energy Dis-aggregation Algorithms like Denoising Autoencoder, RNN, LSTM & some Classic algorithms.
- **Customer Support Chatbot**
  GitHub Link
  - Guide: Dr. Prasenjit Majumdar (Associate Professor at DA-IICT)
  - Conversational bot which solves user queries using sequence to sequence models like LSTM.
  - Understood basic NLP and IR Techniques.

Publications

- **Intent Detection and Discovery from User Logs via Deep Semi-Supervised Contrastive Clustering**, NAACL 2022 (Main Track).
  Paper Link
- **Towards reproducible state-of-the-art energy disaggregation**. In Proceedings of the 6th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation (BuildSys '19). ACM, New York, NY, USA, 193–202.
  Paper Link
- **"Image based Indian monument recognition using convoluted neural networks"** 2017 International Conference on Big Data, IoT and Data Science (BID), Pune
  Paper Link
- **"Evaluating Machine Learning Models for Disparate Computer Systems Performance Prediction"** 2020 IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT)
  Paper Link
- **"Cross-Platform Performance Prediction with Transfer Learning using Machine Learning"** 2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT)
  Paper Link
- **Book Chapter (Springer) - "Modeling Performance and Power on Disparate Platforms using Transfer Learning with Machine Learning Models"** International Conference on Modeling, Simulation and Optimization CoMSO 2020
  Paper Link

Recent Reviews

- 18th International Conference on Natural Language Processing (ICON 2021)
- Asian Journal of Probability and Statistics (ISSN- 2582-0230)
- The Eleventh International Conference on Smart Grids, Green Communications and IT Energy-aware Technologies (Energy 2021 IARIA)

Position of Responsibility

- Teaching Assistant, DA-IICT ; Subject Taught: Algorithms
- Mentor at Raman Classes, Roorkee ; Addressed queries related to Gate Subject

Achievements

- GATE EXAM AIR (2017): 3,301 ; JEE MAINS AIR (2013): 24,236
- Merit Certificates and Gold Medals in X and XII.