Sagarnil Das

sagarnildass@gmail.com • +918420117409
LinkedIn • GitHub • Kaggle • Website

Senior Machine Learning Engineer

Machine Learning & Deep Learning Engineer with 12 years of experience | Intel Edge AI Scholarship Winner | Ex Udacity Machine Learning Mentor | Ex NASA Researcher | Kaggle Notebook Expert

Meet Sagarnil Das, the AI guy who's spent over 12 years making machines smarter and lives better. Whether he's partnering with global giants like the NHS, or casually building a mental health app that detects emotions and suicidal tendencies (because why not?), Sagarnil is all about blending Machine Learning and Deep Learning with humanity.

His journey? A mix of late-night coding, big-picture thinking, and an unrelenting drive to create Machine Learning and Deep Learning solutions that actually matter. When he's not strategizing technological breakthroughs at work, he delves into cutting-edge AI research, and explore the intersections of quantum mechanics and general relativity.

Areas of Expertise

- Generative AI
- Software Development
- Systems Architecture
- Survival Analysis
- Deep Learning
- Reinforcement Learning
- Model Optimization
- Team Building & Leadership
- Cloud Infrastructure Management

Key Accomplishments

- **Architected and Scaled AI Ecosystem:**
  - Achieved significant increases in product adoption and client retention by leading the development of the ARTELUS Technology Ecosystem, steering scalable AI model deployment on cloud platforms like AWS and Azure.
  - Spearheaded the design of Kubernetes clusters and integrated microservices, leading to enhanced clinical workflows.
- **Innovated LLM-based Mental Health Diagnostics:**
  - Conceptualised LLM-based mental health app, incorporating an AI-enabled Patient-360 engine for holistic diagnosis, including emotion detection and suicidal tendency analysis.
  - Engineered a solution that runs 8 LLM models in parallel and in real-time, significantly enhancing the diagnostic capabilities of healthcare providers.
- **Partnership with NHS for AI Retinal Image Analysis Study**:
  - Attained impressive sensitivity rates of 98.0% for moderate-to-severe non-proliferative DR and 98.7% for proliferative DR with the DRISTi 2.1 algorithm.
  - Partnered with the NHS in a landmark study to evaluate AI retinal image analysis systems (ARIAS) for diabetic retinopathy detection.
  - Validated and refined AI technology through this collaboration, paving the way for broader clinical adoption.

Career Experience

Artelus | Kolkata **2023 – Present**
Director and Chief Technology Officer

Architect ARTELUS Technology Ecosystem, pioneering a suite of home-grown tools for scalable AI model deployment. Engineer Artelus cloud ecosystem in AWS and Azure, integrating Kubernetes clusters, RDS and NoSQL databases, and micro-services. Orchestrate training and deployment of deep learning models (EfficientNet, ConvNext) for multi-label classification of retinal conditions, encompassing DRISTi integration. Formulate FDA filing documentation, detailing technical specifications, model architectures, pipelines, and architectural diagrams.

- Strategic Leadership in AI and Healthcare Innovation: Spearheaded the architectural design and strategic roadmap of the Artelus Technology Ecosystem, creating an integrated, scalable infrastructure for AI-

driven healthcare solutions. Defined long-term product vision and technology strategy to position Artelus as an industry leader.

- Created a framework combining natural data and unlabelled fundus images for unsupervised pre-training of robust medical foundational models using SimCLR and BYOL, enabling data-efficient generalisation across various medical tasks.
- International Clinical Study and NHS Collaboration: Led Artelus' engagement in a groundbreaking clinical study with the NHS Diabetic Eye Screening Programme, funded by the NHS Transformation Directorate and The Health Foundation, facilitated by NIHR. The study, titled Ethnic differences in performance and perceptions of Artificial Intelligence retinal image analysis systems for the detection of diabetic retinopathy, showcased DRISTi 2.1's exceptional sensitivity, achieving 98.0% for moderate-to-severe non-proliferative DR and 98.7% for proliferative DR, with results stable across age, sex, ethnicity, and socioeconomic subgroups. This collaboration highlighted DRISTi's potential for large-scale deployment within NHS programs, underscoring its value in diverse healthcare environments and reinforcing Artelus' position as a trusted partner in clinical AI advancements.
- Currently leading the statistical analysis, validation processes, and model verification for FDA submissions, overseeing documentation and compliance for regulatory approvals.
- Cloud Infrastructure and Operational Excellence: Architected Artelus' cloud infrastructure on AWS and Azure, including Kubernetes clusters, RDS and NoSQL databases, and robust microservices. Established a resilient, scalable environment to support high-performance AI operations across platforms.
- Advanced AI Model Development and Deployment: Directed the deployment of advanced models (EfficientNet, ConvNext) for multi-label retinal classification, including integration of the DRISTi platform, delivering real-time insights to healthcare providers.
- Innovation in Synthetic Data and Research Publications: Initiated and supervised the training of StyleGAN3 for synthetic data generation, contributing to enhanced model training datasets. Guided research publication processes, including submissions to high-impact journals such as the British Medical Journal (BMJ).
- Product Development in SaaS and AI-Driven Diagnostics: Conceptualized and directed the development of the Inference Engine SaaS app for automated, point-and-click model deployment. The platform integrates seamlessly with Kubernetes, accelerating time-to-deployment for new healthcare solutions.
- Patient-Centric AI Solutions: Launched the development of a mental health AI engine for holistic patient diagnosis, including emotion, intent, sentiment analysis, and risk assessment. Drove innovation in patient-centric AI with the DRISTi central server, enhancing provider access to comprehensive screening data.

CyberDeck | Kolkata **2022 – 2023**
Director and Chief Executive Officer

Pioneered a no-code platform for end-to-end Data Science and Machine Learning, enabling one-click execution of Data Processing, Exploratory Data Analysis, Inferential Statistical Tests, Machine Learning, Time-Series Forecasting, Clustering, and MLOps. Steered cloud deployment and management through a microservices architecture within a Kubernetes cluster on AWS. Configured Database structures, Load Balancers, Auto Scaling, Grafana, and ELK stack for robust monitoring. Employed S3 and RDS for efficient storage solutions, while integrating numerous Lambda functions for diverse in-app functionalities.

- Cultivated a user base of approximately 5K individuals for CyberDeck's free tier, demonstrating product-market fit and user engagement.
- Developed the entire backend infrastructure of the platform, leveraging cutting-edge technologies and best practices in software development.
- Attained acceptance into Microsoft Startup Founders Hub, garnering $25K in Azure cloud credits to further platform development and scalability.
- Secured selection in Startup-India, a prestigious initiative by the Prime Minister of India, recognising CyberDeck among the most promising startups in the country.
- Qualified for AWS Activate program, acquiring $10K in AWS credits to enhance cloud infrastructure and services.

Hopscotch | Bengaluru **2020 – 2022**
Lead Data Scientist

- Developed a purchase/demographics-based user segmentation model with an unsupervised ML algorithm (Gaussian Mixture model) for better user targeting.
- Improved CTR, RPI, and UPI by creating an optimisation model for Product Listing Page Sort, utilising a tree-based ML model to predict a weighted sum of CTR, RPI, and UPI, and incorporating SHAP values for feature weighting in the ranking framework

- Increased total margin by 4% within 2 months of implementing a Dynamic Pricing model for 10 product types, with individual margins rising by 30% to 172%, using Bayesian Optimization and Markov chain Monte Carlo Simulation.
- Strengthened customer LTV by building a customer survival analysis with Kaplan Meier and Cox's proportional Hazard models, outperforming regression models with censored data.
- Boosted PDP Click Through Rate from 1.3% to 3.2% by introducing multiple recommender systems for Product to Product (Word2Vec and BERT embedding method) and Customer to Product (ALS, Neural Networks, Singular Value Decomposition).

Tathastu(Future Group) | Kolkata **2017 – 2020**
Manager, Data Science Team

- Developed a deep learning model (YOLO/Faster RCNN) to predict age, gender, and emotion from live video feeds, with ongoing use case development. Improved customer profiling with an LSTM model to predict gender, mother tongue, religion, and community from names, contributing to the Member 360 Framework. Streamlined SKU-level pricing using a dynamic price optimisation method with Bayesian optimisation and Markov Chain Monte Carlo (MCMC).
- Slashed campaign operation costs by 4% with a Markov chain Monte Carlo model and Erlang C implementation for queue optimisation in Big Bazaar, paired with an Encoder-Decoder deep learning architecture and Monte Carlo-based error correction for customer visit prediction.
- Increased margin by 7% by predicting customer spending on the next visit using a supervised machine learning model (Xgboost).
- Intensified campaign management with an unsupervised machine learning model (Gaussian Mixture model with PCA) for customer segmentation based on purchase patterns and demographics.
- Decreased operation costs by 4% through a deep neural product embedding model combined with visit prediction to identify potential customers for product promotion campaigns.

Edupristine | Kolkata **2017**
Big Data and Data Science Faculty

Facilitated interactive sessions, demonstrations, and real-life project-solving in domains such as banking and social media. Utilised technologies including Python, Sklearn, Hadoop, Pig, Spark, Hive, and HBase.

- Delivered engaging instruction for the Big Data and Data Science Certification Course, managing classes of 40-50 students.

New York State Department of Health | Albany **2015 – 2017**
Senior Data Specialist | 2016 – 2017

Clustered medical facilities using unsupervised learning models like KMeans and Gaussian Mixture Models based on location data. Managed data migration from Oracle to Hadoop, creating Shell Scripts for data transfer and setting up and working with Hadoop ecosystem tools such as Pig, Hive, Impala, Sqoop, and Spark.

- Spearheaded healthcare facility clustering using advanced unsupervised learning models such as KMeans and Gaussian Mixture Models, optimizing service delivery based on location data and accessibility metrics.
- Led a large-scale data migration project, transferring critical healthcare datasets from Oracle to Hadoop. Designed and implemented Shell Scripts to streamline data transfer processes, achieving a 30% reduction in data retrieval times.
- Supported the department’s Medicaid and Affordable Care Act initiatives by developing predictive analytics solutions to identify high-risk populations and optimize resource allocation, leading to a 5% improvement in healthcare service delivery for underserved communities.
- Delivered comprehensive performance metrics on healthcare program implementations, using Spark 1.6 and SparkSQL to generate weekly analytics reports that drove strategic decision-making at the executive level.

Junior Data Specialist | 2015 – 2016

- Implemented machine learning models for anomaly detection in healthcare claim data, reducing rejection rates by 5% and lowering defect incidence by 20%, thereby increasing overall operational efficiency.
- Conducted extensive correspondence analysis and contingency planning to improve patient data accuracy and reduce misclassification errors by 15%.
- Developed data-driven solutions to optimize reporting processes, enabling a 25% faster turnaround time for critical Medicaid compliance reports.

- Played a pivotal role in transitioning legacy systems to modern data ecosystems, ensuring alignment with federal healthcare regulations and enhancing scalability for future demands.

Computer Science Corporation | Albany **2013 – 2014**
Database Developer

- Contributed to the development of systems for the Affordable Care Act (ObamaCare), which provided healthcare services to 32 million Americans.

Education

Master of Science in Engineering – University of Central Florida | Orlando | 2013
Received the prestigious "Dean's Fellowship" | 2011

- Investigated strain evolution in TBC system through depth measurements using in-situ transmission synchrotron X-ray diffraction.
- Analysed measured strain values to identify critical stages of strain evolution within the TGO.
- Contributed to a NASA-funded SBIR project on CMAS and overlay coatings (La2Zr2O7).

BE, Metallurgical and Materials Science – Jadavpur University | 2011

Technical Proficiencies

Machine Learning | Amazon Web Services | Application Programming | Interface (API) | Artificial Intelligence | Data Science | Kubernetes | Reinforcement Learning | Deep Learning | Robotics | Hadoop | PySpark

Professional Development & Certification

- **Generative Adversarial Networks Specialization**, Coursera | February 2024
- **Deep Reinforcement Learning Nanodegree**, Udacity | December 2022
- **Flying Car and Autonomous Flight Engineer Nanodegree**, Udacity | April 2021
- **Cloud Devops Engineer Nanodegree**, Udacity | March 2020
- **Machine Learning with Python-From Linear Models to Deep Learning**, MIT-EdX | September 2019
- **Probability - The Science of Uncertainty and Data**, MIT-EdX | January 2019
- **Robotics Software Engineering Nanodegree**, Udacity | October 2018
- **Deep Learning Specialization**, Coursera | March 2018
- **Artificial Intelligence Nanodegree**, Udacity | March 2018
- **Machine Learning Engineer Nanodegree**, Udacity | August 2017

Publications

AI-Driven Diabetic Retinopathy Screening: Multicentric Validation of AIDRSS in India | January 2025
arXiv Electrical Engineering and Systems Science - Image and Video Processing (arXiv:2501.05826)

Enhancing Early Diabetic Retinopathy Detection through Synthetic DR1 Image Generation: A StyleGAN3 Approach | January 2025
arXiv Electrical Engineering and Systems Science - Image and Video Processing (arXiv:2501.00954)

Calculating Customer Lifetime Value and Churn using Beta Geometric Negative Binomial and Gamma-Gamma Distribution in a NFT based setting | January 2025
Social Science Research Network (10.2139/ssrn.5087590)

Robot Localization in a Mapped Environment Using Adaptive Monte Carlo Algorithm | October 2018
International Journal of Scientific and Engineering Research (Vol 9, Issue 10, Oct 18, 2018 – ISSN 2229 5518)

Simultaneous Localization and Mapping (SLAM) using RTAB-Map | August 2018
International Journal of Scientific and Engineering Research (Vol 9, Issue 8, August 2018 – ISSN 2229 5518)

Speaking Engagement

- **Keynote Speaker at Jadavpur University | March 2019**
  Invited as Keynote speaker at Jadavpur University – Artificial Intelligence Conference