Sayan Dasgupta

(+91)9591374000 • dasgupta.sayan@gmail.com • sayandasgupta

EXPERIENCE

Citrix, Netscaler Business Unit

Principal Data Scientist — *Dec 2020 – Present, BLR, India*

- Architected and developed ***Netscaler Copilot***, a RAG-based AI platform enabling question-answering across Netscaler products, reducing time-to-resolution (TTR) for customers by 20%.
- Designed and implemented guardrails to protect the system from adversarial attacks.
- Created evaluation metrics (BERT Score and BLEU Score) to measure the AI model's performance.
- Developed models to identify volumetric and sophisticated **Account Takeover attacks**, leveraging statistical and machine learning techniques.
- Built sentiment analysis models to analyze customer feedback and identify key areas for improvement.
- Developed the ***Netscaler Voice of Customer (VoC) Dashboard*** to visualize insights and track progress on improvements.

Reliance JIO Embibe

Principal Data Scientist — *Apr 2019 – Dec 2020, BLR, India*

- Led a 10-member data science team to develop an AI-based learning platform.
- Built a learning journey recommendation system, increasing user engagement from 16 minutes to 23 minutes (**43%**) and improving learning outcomes by 20%.
- Developed and deployed ***Bayesian Knowledge Tracing (BKT)*** models to measure student mastery of various concepts.
- Built **2PL IRT** models to understand question difficulty and discrimination for test design.
- Conducted A/B testing to validate the impact of the models on the platform.

Careem

MTS in Data Science — *Apr 2018 – Mar 2019, Dubai, UAE*

- Developed Careem's ML-based driver fraud detection system, significantly increasing cost savings.
- Created statistical and ML models to identify phantom trips.
- Built a system to augment automated trip-flagging with manual review by local teams, enabling timely intervention.

C1X Inc

Senior Manager of Data Science — *Oct 2016 – Sept 2017, BLR, India*

- Implemented real-time bidding (RTB) enrichment via extensive A/B testing on ad impression bid prices.
- Architected a data science and analytics platform using AWS, EMR (ETL), MySQL, and Apache Superset (BI). Reduced time to insights from multiple days to under one hour.

LinkedIn

Senior Data Scientist — *May 2013 – Sept 2016, BLR, India*

- Designed and implemented machine learning models to detect fake accounts, preventing fraud and spam. Reduced touch back rate to **1%** from north of 4%.
- Reduced invitation spam by **18%** over control group leveraging the member identity model.
- Built end-to-end model training pipelines, handled feature engineering, and led model evaluation.
- Trained and deployed classifiers to detect spam in short text fields such as user names and headlines.

Shopzilla, Symphony

Senior Data Scientist — *Jan 2012 – Apr 2013, BLR, India*

- Forecasted Revenue Per Click for SEM keywords using big data technologies and statistical/machine-learning models.

Innovation Labs, 24/7 Inc

Senior Analytics Consultant — *Aug 2010 – Jan 2012, BLR, India*

- Built and deployed machine learning models utilizing clickstream data to score visitor propensity to churn or purchase.

EDUCATION

2007 – 2009
Master of Statistics
Indian Statistical Institute, Delhi

2004 – 2007
B.Sc. in Statistics (Honours)
St. Xavier's College, Kolkata

Invited Talks

2015
Facebook Spam at Scale
Presented research on detecting spam in short text. Link

2016
Smart Data Summit, Dubai
Discussed leveraging data science to prevent abuse on LinkedIn.

Achievements

- Secured **All India Rank 15** in IIT JAM (Statistics) 2007.
- Patent No: US10204307B1
- Received the prestigious **Rajiv Sinha Technology Award** at Netscaler.

Technical Skills

- **Programming Languages:** Python, R, SQL, Bash, LaTeX
- **Machine Learning:** scikit-learn, Prophet, XGBoost, Statsmodels
- **Deep Learning:** PyTorch
- **Generative AI:**
  - **LLM Frameworks:** LangChain, Autogen, FAISS, Qdrant
  - **Models & APIs:** Azure OpenAI, Anthropic Claude
- **Big Data:** Spark, Hadoop, Hive
- **Cloud Platforms:** Azure, AWS

Selected Technical Writeups

- Backpropagation in PyTorch
- Logistic Regression

References

Available upon request.