Avinash Kumar (アヴィナシュ.クマル)

Data Scientist - Rakuten, Ex-AstraZeneca

• **MBA - Hitotsubashi University** • **(+81) 90-2389-7660** • **(+91) 784-585-5206**

• avinashkr1992@gmail.com • www.linkedin.com/in/avinash1992

Summary

Overall: Results-Driven Data Scientist | 9 Years of Expertise in Pharmaceuticals, Telecom, and Consultancy | Proven Leadership and Communication Skills Generating Millions in Revenue | Storytelling | Specialized in End-to-End Execution of Data Science Projects | Mentor to Emerging Data Scientists | Innovation | Reliable | Leadership | Strategic thinking | Cross-functional collaboration | MEXT YLP Scholar in Japan

Tech.stack:
- Python (Data collection, Data wrangling, EDA, Model development, Deployment, Monitoring Automation)
- Machine Learning, Deep Learning, Statistics
- Natural Language Processing
- SQL & Git
- Linux, Unix, Bash Scripting
- Kubernetes, Dockers, Argo workflow.
- SwiftUI iOS App Dev
- AWS, Azure, GCP
- Knime, Dataiku, H2O, R, Tableau
- LLM Tuning & Web UI setup
- Orange, Glueviz, SIMCA, SAS

iOS Apps.
- Planet Spotter App: An app that shows the live location of the celestial bodies
- Voice Clock & ClockFace App **:** An accessibility-focused clock app with Voice
- Hikers' View: Real-time landmark tracking app

Languages: English (Native), Japanese (N3, actively improving towards N2), Hindi (Native), Tamil (Conversational)

Experience

Rakuten Group
Tokyo, Japan

Data Scientist, Measurement & Insight Section
Jan 2023 – Present

- **Advanced Customer Analytics and ML-Driven Acquisition Strategy for Rakuten Insurance:** Architected end-to-end machine learning pipeline combining K-means clustering for customer segmentation (50+ features) with lookalike modeling to identify high-potential prospects. Engineered feature space spanning demographics, transaction patterns, and digital footprints to build similarity scoring algorithm for targeted advertising. Implemented A/B testing framework, achieving 1.5% increase in sales conversion across insurance products (life, medical, car, cycle, pet).

- **ML Pipeline Orchestration and Automated Analytics:** Architected end-to-end machine learning pipelines in Hive for real-time feature engineering and model serving, enabling automated customer scoring and segment assignment. Implemented automated ETL processes with data quality monitoring, AI governance, ethics, automation and validation checks and ensuring reliable training data for predictive models

- **AI-Powered Analytics and Insight Generation Platform - Implemented Generative AI solutions for automated decision-making and knowledge extraction** - Architected end-to-end ML pipeline combining NLP to automatically generate comprehensive sales reports. Implemented custom LLM fine-tuning for domain-specific analysis, with automated data validation. Developed real-time monitoring system for key performance metrics and trend detection, reducing report generation time by 8 hours while achieving 95% accuracy.

AstraZeneca
Chennai

Data Scientist, Operations & IT
Oct 2019 – Sep 2021

- **Dispense Planning Optimization - Optimized pharma dispense planning, integrating AI models with GMP regulations, achieving $330K savings:** Developed a mathematical optimization model using IBM CPLEX with mixed- integer linear programming to streamline material dispensation in production. The model, aimed at reducing human effort.

- **Developed AI-driven complaint monitoring system improving regulatory compliance and risk mitigation:** Enhanced complaint monitoring for regional insights through automated analysis of GMSGQ data, employing AI and Natural Language Processing. Utilized TD/IDF models with AWS Sagemaker, Presto & S3 for deployment, resulting in a daily time saving 16 man-hours and an improved and expedited review process for complaints.

Efficiency Projects:

⇒ **Digital Line Clearance:** Utilized CNN Deep learning to provide real-time status updates online clearance between the production of two batches. Successfully reduced the clearance time from 1 hour to just 7 minutes.

⇒ **TrackSys Data Analytics:** Conducted extensive data mining to optimize the production line and Overall Equipment Effectiveness (OEE) via Predictive Maintenance Modeling and Down Time Analysis, leveraging GMSGQ data and provided actionable recommendations for business or content improvement

⇒ **Carbon footprint Prediction:** Built dashboards for visualizations to self-serve metrics and trends effectively and implemented an IoT system on Quicksight utilizing logistic regression and other statistical models, mathematical modelling and forecasting to monitoring of carbon emissions.

⇒ **Repeat Deviation Analysis:** Enhanced deviation resolution time globally by implementing similarity scoring. This resulted in a 30% reduction in ticket duplication per month, leading to faster and more efficient results.

- **Breztri Inhaler Optimization** Utilized descriptive statistics for Proof of Concept (PoC) to examine reduced shot weights of medication, aiming to enhance the consistency in the amount of medicine delivered in an inhaler shot.
- **Developed AI governance framework ensuring ethical AI practices and regulatory compliance – URS** Collaborated and authored a comprehensive User Requirement Specification (URS), along with other business stakeholders, which was widely adopted to comprehend user needs and preferences. Influenced cross-functional stakeholders, aligning AI initiatives with global business objectives.

Tata Consultancy Services **Chennai, Tokyo**
Data Scientist, IoT, TEPCO ***Dec 2014 – Oct 2018***

Led a dynamic team of 7 in an Industrial IoT project focused on optimizing power plant operations and reducing pollution. Spearheaded Data Analytics and delivered Artificial Intelligence (AI) projects.

- **Intelligent Power Plant Optimization: Digital Twin and Performance Enhancement**: Directed a team in executing Intelligent Power Plant projects, including Digital Twin Modeling and Performance Optimization for vendors such as TEPCO, KEPCO, and Tata Power. Employed Particle Swarm Optimization, dimensionality reduction and PLS regression to achieve maximum efficiency and minimize emissions

- **Successfully reduced pollution discharge at TEPCO power plants:** Successfully reduced pollution discharge at TEPCO coal-fired power plants by 10% through strategic adjustments in furnace settings. Implemented cost-cutting measures, resulting in a notable $10 million in savings, recognized and featured in Nikkei Asia. (https://asia.nikkei.com/Business/Tata-Japanese-partners-to-use-AI-at-coal-fired-power-plants).

Education

Hitotsubashi University Business School, School of International Corporate Strategy (ICS) ***Tokyo, Japan***
Master of Business Administration Candidate ***(MEXT YLP Scholar)*** ***20 Aug 2022***

Anna University ***Chennai, India***
Bachelor of Engineering in Electronics and Instrumentation ***31 May 2014***

- Awarded 'Best Outgoing Student' for 2010-2014
- Published a technical paper "Mango-meter" in National Level Conference

---

Awards & Projects

- Built new Alexa Skill in Hackathon jointly conducted by AstraZeneca and AWS (EC2, Lambda), commands for automatically counting inventory and all existing product in a warehouse (placed 3rd in the event)
- Ranked 9 in TechGiG Sales prediction ML Challenge. https://www.techgig.com/challenge/machine-learning-1
- 3 Stage dance performances viz Bhangra & Runner up British Telecom - TCS Chess Championship 2016
- Won 6 Paper presentations and 19 technical and non-technical Awards in National Level Technical
- Awarded 'Best Outgoing Student' for 2010-2014 of the department in the Bachelors'.