

KAMESH DUBEY

Mumbai, India

+91-7217563163

kameshcodes@outlook.com

LinkedIn

Portfolio

PROFESSIONAL SUMMARY

Data Scientist applying statistical and machine learning methods to build risk solution. Proficient in Python, SQL, and the ML model development lifecycle.

TECHNICAL SKILLS

Technical Skills: Python (Pandas, Scikit-learn, PyTorch, LangChain), R, Excel, SQL

Key Coursework: Machine Learning, Statistical Inference, Time Series Forecasting

WORK EXPERIENCE

Associate Consultant – Financial Risk Management

Jun 2025 – Present

KPMG Assurance and Consulting Services LLP

Mumbai, India

Model Risk Validation (Client Engagements)

- Performed quantitative validation of PD scorecards using **KS statistic**, **Hosmer–Lemeshow**, and **PSI** metrics
- Conducted independent validation of **Value-at-Risk (VaR)** and **Bond Valuation** models through **back-testing**, **stress testing**, and **sensitivity analysis**, supporting forecasting and capital risk assessments.

Graph RAG Chatbot (Global Initiative, KPMG Germany)

- Integrated a **re-ranking module** on the top of extracted document content improving **nDCG@k** and **Recall@k**
- Implemented an **OCR** pipeline that **delivers 17% F1** gains with **CPU-only inference** enabled by quantization

Data Science Intern

Apr–May ‘25

Bluepond AI

Chennai, IN

SOV Chatbot for Underwriters

- Built a RAG chatbot with tool-calling capabilities to generate visualizations and insights from the Excel files
- Designed an **AI-driven data pipeline** to extract, clean, and standardize Statement-of-Value (SoV) Excel files
- Fine-tuned a **BERT (99.5% F1)** classifier on a custom dataset to filter non-insurance queries before API calls

Brokers Co-Pilot

- Built an **extraction tool** that extracted **30+** key fields from insurance documents, cutting review time by **80%**
- Revamped the document reader pipeline to enhance text parsing precision and improve extraction accuracy

KEY PROJECTS

Time Series Analysis | Course Project

Jan–Apr ‘24

Guide – Sanjeev V Sabnis, Department of Mathematics, IIT Bombay

- Conducted a thorough **statistical analysis** of daily sales data and applied **ADF test** to assess stationarity
- Performed **Ljung-Box test** on the ARIMA model, **backtested** it, and forecasted the next 6 months of sales

Bank Customers Segmentation | Course Project | GitHub

Aug–Nov ‘23

Guide – Prof. Siuli Mukhopadhyay, Department of Mathematics, IIT Bombay

- Enhanced interpretability by leveraging **PCA** and **factor analysis** to uncover latent variable in the dataset
- Segmented data using **K-Means** clustering and evaluated clusters via **Silhouette** Score and Elbow Method

EDUCATION

Indian Institute of Technology, BOMBAY

Aug’22 - Dec’24

M.Sc. Applied Statistics and Informatics

Mumbai, IN

Gurukula Kangri University

Aug’18 - Jul’21

B.Sc. Mathematics and Computing

Haridwar, IN

SCHOLASTIC ACHIEVEMENTS

- Secured **AIR-50** in Mathematical Statistics, IITJAM 2022 among 2,912 candidates organized by IITR ‘22
- Achieved a remarkable **Department Rank 3** out of 120 BSc. Students at Gurukul Kangri University ‘21