

KOUSHIK KHAN

+91 (750) 192-3251 Bangalore, India

koushik.khan.ds@gmail.com linkedin.com/in/koushikkhan/ koushikkhan.github.io

SUMMARY

Full-Stack Data Scientist with 10 years of experience in delivering data-driven solutions across diverse domains. Skilled in end-to-end data science workflows - from data sourcing and ETL to analytics, ML/DL modeling, and reporting. Experienced in driving cross-functional analytics initiatives that support organization-wide goals.

SKILLS

| | |
|-------------------|---|
| Data Science | Data Analysis, Forecasting, Predictive Modeling, NLP, Recommender Systems |
| Tech Stack | Python, R, SQL, PySpark, PyTorch, TensorFlow |
| Platforms & Tools | Linux, Containerization, CI/CD, AWS, Databricks |
| Domain Expertise | Supply Chain Analytics, HR Analytics, Product Development |

WORK EXPERIENCE

Lead Data Scientist

October 2024 - Till Date

Tiger Analytics India Consulting Pvt. Ltd.

Bangalore, India

- Leading a team to design, develop, and maintain in-house data science frameworks and tools, collectively known as **TigerML**, that serve as the backbone for critical customer projects. These frameworks standardize workflows, optimize resource usage, and ensure scalability, enabling faster data analysis, model development and deployment. As a result, delivery timelines for customer-facing solutions have been accelerated by **60%**, enhancing productivity, consistency, and overall business impact.
- Leading analytics initiatives for the HR team, focusing on building tools that expedite analyzing HR data and automate reporting workflows, develop dashboards and organization-wide reports to track the performance and impact of in-house technical training programs, enabling data-driven decision-making.
- Actively contributing to multiple customer-facing projects, driving solution development from initial discovery to prototype and team setup. Collaborating closely with stakeholders to define requirements, design PoCs, and build scalable data-driven solutions. Recent engagements include developing a RAG application for a banking client and a pricing recommendation app for a retail client.

Skills Used: Python, PySpark, PowerBI, Linux, Deep Learning, LLM, LangChain, Docker, AWS

Principal Data Scientist

September 2023 - September 2024

S&P Global, Market Intelligence

Bangalore, India

- Led a cross-functional team in designing and developing advanced predictive models and robust DevOps pipelines for handling large bill-of-lading datasets for the maritime supply chain and emission products. Historical data, analytics and machine learning techniques had been leveraged to forecast supply chain dynamics and ship emissions supporting strategic decision-making for the maritime business.
- Engineered a NLP application to automatically detect potentially risky bill-of-lading documents by identifying dual-use goods (having military and commercial usages). By applying advanced NLP techniques on top of large open source foundation models (LLMs), the application efficiently flagged high-risk documents, reducing the likelihood of non-compliance and trade financing risks. The application contributed to almost **13% improvement** in the financial decision-making process, allowing for faster approvals and more secure trade transactions.

Skills Used: Python, PySpark, Deep Learning, LLM, DevOps, AWS

Senior Data Scientist

S&P Global, Commodity Insights

August 2021 - September 2023

Bangalore, India

- Led the cloud migration for a major ETL codebase of the maritime analytics product ([Commodities At Sea](#)), improving scalability and reducing manual effort by **40%**. Developed modular Python packages and standardized workflows, enhancing maintainability and cutting deployment time. This shift improved operational efficiency and contributed to a **15% revenue increase** through new client acquisition.

Skills Used: Python, PySpark, CI/CD, Linux, DevOps, AWS

Data Scientist

IBM India Pvt. Ltd.

October 2016 - July 2021

Bangalore, India

- Served as a Machine Learning Engineer for a leading automobile client, developing an automated vehicle diagnostic tool using a fine-tuned BERT model on transcribed call logs. The solution **reduced false warranty claims by 14%**, cutting operational costs. Also led model productionization, ensuring seamless integration and scalable deployment.
- Drove an average **10% revenue growth** across multiple markets by developing market-mix modeling (MMX) to optimize marketing strategies for a leading retail client. Focused on beverage sales in Indonesia and Pakistan, enabling more effective resource allocation. Developed automated frameworks for reporting and visualization, delivering real-time, actionable insights that improved decision-making and reporting efficiency.
- Developed a contextual FAQ search engine using Latent Semantic Indexing (LSI), improving search relevance and **boosting customer satisfaction by 20%**. The solution captured semantic relationships beyond keyword matching, enhancing user experience and engagement. This success secured **\$50K** in funding and unlocked new business opportunities with one of the leading UK based banking clients.
- Developed an NLP and deep learning-based framework to automate tech support ticket analysis and root cause identification as part of [IBM Ignite](#). Leveraged LSTM with Attention to achieve high classification accuracy, streamlining support workflows and enabling faster issue resolution. Successfully deployed across banking and insurance clients, securing a **\$200K deal** and funding for further development while improving customer satisfaction and reducing operational costs.

Skills Used: Deep learning, TensorFlow, Docker, Kubernetes, RESTful APIs

Data Analyst

BCausE Enterprise Pvt. Ltd.

October 2015 - August 2016

Noida, India




- Built automation tools to scrape customer reviews for sentiment analysis, eliminating manual data collection and accelerating feedback cycles. Developed interactive dashboards that streamlined monthly reporting, **reducing manual effort by 20%** and enabling faster, data-driven decisions.

Skills Used: Python, MongoDB, R Shiny

EDUCATION

- **M.Sc., Statistics:** Visva Bharati University, West Bengal, India 2013 - 2015
- **B.Sc., Statistics:** Visva Bharati University, West Bengal, India 2010 - 2013

CERTIFICATIONS

- [Tiger Analytics Certified: Cloud Deployment Foundation 100](#) 
- [Microsoft Certified: Azure Fundamentals](#) 
- [Microsoft Certified: Azure Data Scientist Associate](#) 
- Microsoft Certified: Azure AI Fundamentals