KOUSHIK KHAN

☎ +91 (750) 192-3251 ◇ 📍 Bangalore, India

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

SUMMARY

Full-Stack Data Scientist with over 10 years of experience in delivering data-driven solutions across diverse domains. Proficient in managing the entire data science lifecycle - from data acquisition and ETL to applying advanced analytics, machine learning, deep learning, and reporting. Experienced in driving cross-functional analytics initiatives that support organization-wide goals.

SKILLS

Data Science	Data Analytics, Predictive Modeling, NLP
Tech Stack	SQL, R, Python, PySpark, PyTorch
Platforms & Tools	AWS, GCP, Databricks, Linux, Containerization, CI/CD, DevOps
Domain Expertise	Supply Chain Analytics, Retail Analytics, Product Development

WORK EXPERIENCE

Lead Data Scientist <div align="right">October 2024 - Till Date</div>
Tiger Analytics India Consulting Pvt. Ltd. _Bangalore, India_

- Leading a cross-functional team to develop price elasticity models and scenario planners for a retail customer on the Indian beauty market. I oversee the end-to-end lifecycle of data processing pipelines on GCP, utilizing Bayesian models and statistical analyses to derive granular insights across diverse product categories. Recently streamlined delivery by implementing automated data validation and integrity frameworks, significantly reducing processing time and ensuring high-fidelity model inputs.

- Led 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.

- Led 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 contributed to multiple customer-facing PoC's, driving solution development from initial discovery to the prototype development and team setup. Collaborated closely with stakeholders to define requirements and build scalable 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, GCP

Principal Data Scientist <div align="right">September 2023 - September 2024</div>
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   August 2021 - September 2023

S&P Global, Commodity Insights   *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   October 2016 - July 2021

IBM India Pvt. Ltd.   *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   October 2015 - August 2016

BCausE Enterprise Pvt. Ltd.   *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