

# Koushik Khan

Bengaluru - 560049, Karnataka, India  
Last Updated: May 17, 2024

+91 7501923251  
koushik.khan.ds@gmail.com  
linkedin.com/in/koushikkhan  
koushikkhan.github.io

## SUMMARY

A seasoned data scientist with a master's degree in Statistics and nearly nine years of comprehensive experience in the data analytics industry. Demonstrated expertise in crafting data-driven products and services by utilizing machine learning and deep learning techniques, spanning across predictive modeling and natural language processing applications. Looking for opportunities to leverage my skills, experience, and leadership abilities in a vibrant professional setting.

## PROFESSIONAL EXPERIENCE

**Principal Data Scientist**  
S&P Global Inc.

August 2021 - Till Date  
Bangalore, India

- Lead analytics team to develop essential DevOps pipelines and predictive models for maritime supply chain products.
- Develop a NLP application to classify possibly risky bill-of-lading documents, reducing seaborne trade financing risk.
- Develop a predictive model for the fuel oil consumption for container ships, giving estimated fuel usage before a voyage.
- Deployed legacy codebase to DevOps, boosting product performance by having new clients and ~ 15% higher revenue.
- Developed and maintained in-house Python libraries & automation tools, reduced manual efforts by 40%.

Methods & tech stacks used: Statistics, Machine Learning, Deep Learning, Python, PowerBI, DevOps, AWS

**Data Scientist**  
IBM India Pvt. Ltd.

October 2016 - July 2021  
Bangalore, India

- Developed a NLP application streamlining tech support ticket analysis, reduced manual efforts by 15%.
- Envisioned, built and deployed a banking FAQ retrieval system that outperformed existing keyword-based models by 20%.
- Implemented a text summarization model using deep learning from customer feedback data, helped in upstream tasks.
- Developed a NLP model to automate vehicle diagnostics from service call logs, reduced false claims at early stage by 10%.
- Developed text-cleaning & modelling pipelines for IBM Watson sentiment analysis API.
- Accomplished an average 10% sales growth, across markets, using Market-Mix-Model derived from sales and ad data.

Methods & tech stacks used: Deep learning for NLP, Machine Learning, Docker, REST API, Python

**Data Analyst**  
BCauseE Enterprise Pvt. Ltd.

October 2015 - August 2016  
Noida, India

- Collaborated on identifying potential consumers for an in-house product by developing a customer segmentation model.
- Developed dashboard with R Shiny to simplify monthly report generation, depicting consumer activities.
- Developed automation tools to scrap data periodically from web pages and making them structured.

Methods & tech stacks used: Clustering, MongoDB, Python, R Shiny

## KEY SKILLS

- **Programming Languages:** Python, R and SQL
- **Tools & Technologies:** PySpark, PyTorch, Power BI, AWS, Docker, REST API, CI/CD, Azure DevOps, Linux
- **Analytics:** Statistical Modeling, Machine Learning, Deep Learning, Natural Language Processing

## EDUCATION

- **M.Sc, Statistics:** (CGPA 8.43), Visva Bharati University, West Bengal, India 2013 - 2015
- **B.Sc, Statistics:** (CGPA 8.26), Visva Bharati University, West Bengal, India 2010 - 2013

## CERTIFICATIONS

- Microsoft Certified: Azure Fundamentals
- Microsoft Certified: Azure Data Scientist Associate
- Microsoft Certified: Azure AI Fundamentals

## PERSONAL PROJECTS

- Authoring technical articles for my blog as well as other platforms.
- All of my past and ongoing projects are listed here.