

# Nadeem — Data Scientist

**Phone:** +91 9743614471    **Email:** [nadeem2796@gmail.com](mailto:nadeem2796@gmail.com)    **LinkedIn:** [linkedin.com/in/nadeem](https://linkedin.com/in/nadeem)

**Location:** Bengaluru, India (open to relocation to UAE, Saudi Arabia, Qatar)

**Nationality:** Indian    **Languages:** English, Hindi, Tamil, Kannada

## Professional Summary

Data Scientist and Machine Learning Engineer with 5 years of experience delivering data-driven solutions for global clients across automotive, retail, FMCG, finance, and healthcare sectors. Skilled in time-series forecasting, predictive modeling, marketing mix modeling & optimization, generative AI, and agentic workflows. Proficient in building scalable pipelines using big data technologies and deploying models on AWS, Azure, and GCP using MLOps best practices. Experienced in communicating complex insights to non-technical stakeholders, leading multicultural teams, and mentoring junior scientists. Open to remote or hybrid roles and ready to relocate to the UAE, Saudi Arabia, or Qatar to contribute to AI-driven digital transformation.

## Professional Experience

### DC Consultant – Data Science, Deloitte, Bengaluru, India | Nov 2024 – Present

- Developed Model Ready Data (MRD) for a major athletic apparel retailer by joining diverse tables, engineering features, and configuring forecasting asserts; implemented multiple time series models and ML algorithms, achieving a 12% improvement in forecasting accuracy (MAPE) over baseline.
- Automated model selection using accuracy metrics (MAPE, SMAPE, model ranks) and ensembles top-performing models; followed by regression-based breath and depth optimization, segmentation, and imputation using NLP techniques.
- Leveraged GenAI to auto-generate insight summaries and perform store-level, cross mapping; deployed pipelines using GCP services to support scaled, production-grade forecasting.
- Worked with a global CPG client to build MRD and perform extensive feature engineering for demand forecasting; collaborated across teams to ensure data consistency and model explainability.
- Led model validation for a leading retail healthcare client by comparing internal promotional models with client-delivered alternatives; evaluated performance across product, category, brand and segment level using snowflake, kubeflow, mostly using SQL and Python. Automated validation workflows with business rule checks, confidence scoring and visual reporting to ensure model reliability and stakeholder alignment.

### Senior Data Scientist – Media Effectiveness, Dentsu, Bengaluru | Mar 2022 – Oct 2024

- Led end-to-end development of Marketing Mix Models (MMM) for major clients across automotive, healthcare, finance, and retail sectors, improving marking ROI through data-driven optimization.
- Created analytical data sets (ADS), conducted exploratory data analysis (EDA) and built regression-based MMM frameworks to quantify media impact.
- Developed and deployed a budget optimization algorithm to allocate spend across media channels using model coefficient, KPIs and client specific constraints.
- Built a scenario simulator tool to visualize campaign outcomes, enabling clients to simulate investment decisions and fine-tune strategy.
- Designed and maintained automated data pipelines for ingesting media, sales, and external signals using Python, SQL and cloud environment AWS (Athena and SageMaker)
- Contributed to the Pricing and Vehicle Incentive Steering initiative by building an end-to-end incentive optimization model for a leading automobile brand. Prepared sales datasets using PySpark and merged channel and offer-level data for holistic analysis, Implemented a four-stage modeling approach: micro-segmentation, elasticity modeling,

passthrough prediction and optimized discount allocation, resulting in measurable uplift in incentive efficiency and vehicle sales.

- Generative AI Video Analytics; Analyzed video creatives, identifying engagement-driving features leading to a 30% increase in video engagement
- Tool & Tech: AWS SageMaker, Athena, EC2, Python, PySpark, SQL, Postman, Advanced Excel

#### **Data Scientist, Soothsayer Analytics (remote) | Feb 2021 – Mar 2022**

- Collaborated with a global team of data scientists to address manufacturing pain points by converting domain-specific problems into machine learning problem statements using P&ID diagrams and process knowledge.
- Developed and optimized Python modules for real-time data preprocessing, improving data reliability and model readiness.
- Built regression and classification models to predict Remaining Useful Life (RUL) of machinery and detect anomalies, supporting predictive maintenance strategies.
- Worked on supply chain demand forecasting across multiple granularities using time-series models, automating workflows through Azure Data bricks and Data Studio.
- Designed a framework to detect duplicate spare parts using clustering and fuzzy logic, enhancing inventory accuracy and warehouse efficiency.
- Applied Unsupervised learning techniques to identify anomalies in sensor data from manufacturing processes.
- Participated in cross-functional collaboration with engineers, domain experts and stakeholders to deliver insights, train end-users and ensure business alignments.

#### **Data Science Consultant, INSOFE (International School of Engineering), Bengaluru | Aug 2020 – Jan 2021**

- Designed predictive models for early detection of manufacturing incidents using Random Forest and Gradient Boosting, delivering one and four-hour advance warnings.
- Built data preprocessing pipelines in Python for near real-time streaming data integrated with HDFS and Hive using tools like Dataiku & RapidMiner.
- Worked closely with process engineers and domain experts to translate operational challenges into actionable ML solutions.
- Delivered proof-of-concept (POC) models and supported training sessions to upskill internal teams on advanced analytics tools and frameworks.

## **Technical Skills**

---

<b>Languages:</b>	Python, SQL, Pyspark
<b>Machine Learning:</b>	Linear/Logistic Regression, Decision Trees, Random Forest, Gradient Boosting, XG-Boost, SVM, KNN, Naïve Bayes, KMeans, DBSCAN
<b>Deep Learning &amp; Generative AI:</b>	CNN, RNN/LSTM, Transformers, Auto-encoders, GANs, GPT/BERT, diffusion models, LLM fine tuning, prompt engineering
<b>Agentic &amp; Reinforcement Learning:</b>	Multiagent systems, Q learning, policy gradient methods
<b>NLP &amp; Computer Vision:</b>	Tokenisation, embeddings, sentiment analysis, summarisation, translation, image classification, object detection
<b>TimeSeries:</b>	ARIMA/SARIMA, Prophet, exponential smoothing, sequence-to-sequence models
<b>Data Engineering &amp; Big Data:</b>	ETL/ELT, Airflow, Apache Spark/PySpark, Hadoop, Hive, Kafka, Delta Lake, Snowflake
<b>Databases:</b>	MySQL, PostgreSQL, SQL Server

**Cloud & MLOps:** AWS (SageMaker, EC2, S3, Glue, Lambda), Azure (Databricks, Data Factory), GCP (BigQuery, Vertex AI), Docker, Kubernetes, Git, Jenkins, MLflow

**Visualisation & BI:** Tableau, PowerBI, Looker, Matplotlib, Seaborn, Plotly

## Education

---

- **Post Graduate Program in Data Science & Machine Learning** – INSOFE, Bengaluru (coursework in machine learning, deep learning, NLP, big data & business analytics) | Carnegie Mellon University
- **Bachelor of Electrical & Electronics Engineering** – Visvesvaraya Technological University, Karnataka

## Certifications

---

- Deep Learning Specialization, Sequence Models, Convolutional Neural Network, Structuring Machine Learning.
- Introduction to Artificial Intelligence (AI) | IBM
- Introduction to Data Science in Python, Applied Machine Learning in Python | University of Michigan
- Exploratory Data Analysis, The Data Scientists Toolbox, R Programming | The John Hopkins University
- Google Cloud Platform Big Data and Machine Learning Fundamentals | Google
- Machine Learning Professional, Data Engineering Professional Certification | RapidMiner

## Achievements & Interests

---

- Top 10 % finish in Kaggle Demand Forecasting Challenge (2021)
- Speaker at regional data science meetups on generative AI and MLOps
- Mentor for aspiring data scientists; contributor to open source projects; exploring ethical AI and fairness