**Name:** Arjun Mehta  
**Phone:** +91 9876543210  
**Email:** arjun.mehta@example.com  
**LinkedIn:** linkedin.com/in/arjunmehta  
**GitHub:** github.com/arjunmehta  
**Location:** Bengaluru, Karnataka, India

**Professional Summary**

Data Scientist with **5 years** of experience in leveraging data analytics, machine learning, and statistical modeling to drive business insights and decision-making. Proven expertise in handling large-scale data, developing predictive models, and optimizing algorithms for industries such as **Finance, Retail, Healthcare, and Supply Chain.** Skilled in Python, R, SQL, Spark, and cloud platforms like AWS/GCP/Azure. Passionate about deploying AI/ML solutions at scale to solve real-world business problems.

**Technical Skills**

* **Programming:** Python (Pandas, NumPy, Scikit-learn), R, SQL, Spark (PySpark)
* **Machine Learning & AI:** Regression, Classification, Clustering, NLP, Time Series Analysis, Deep Learning (TensorFlow, PyTorch)
* **Big Data & Cloud:** Apache Spark, Databricks, Hadoop, AWS (S3, EC2, SageMaker), GCP (BigQuery, Vertex AI)
* **Databases:** SQL (PostgreSQL, MySQL), NoSQL (MongoDB, Cassandra)
* **Data Visualization:** Tableau, Power BI, Matplotlib, Seaborn
* **MLOps & Deployment:** Docker, Kubernetes, CI/CD Pipelines, Flask, FastAPI

**Work Experience**

**Senior Data Scientist | Infosys | Jan 2022 – Present**

**Industry: Fintech | Risk & Fraud Analytics**

* Developed a **real-time fraud detection system** using **Kafka, Spark Streaming, and ML algorithms** that reduced fraudulent transactions by **35%**.
* Built **anomaly detection models** for payment transactions using **unsupervised learning (Isolation Forest, Autoencoders)** to flag suspicious activities.
* Designed and deployed **credit risk scoring models** leveraging XGBoost and Random Forest, improving loan approval accuracy by **20%**.
* Implemented **automated feature engineering pipelines** in Databricks, reducing model development time by **40%**.
* Deployed models using **AWS Lambda + API Gateway** for real-time scoring, reducing API latency by **30%**.

**Data Scientist | Flipkart | Jul 2019 – Dec 2021**

**Industry: Retail & E-commerce | Customer Behavior Analytics**

* Designed **personalized recommendation systems** using collaborative filtering (ALS) and deep learning (Neural CF), increasing user engagement by **22%**.
* Developed **demand forecasting models** using **ARIMA, LSTM, and Prophet** to optimize inventory, reducing stock-outs by **15%**.
* Conducted **A/B testing and uplift modeling**, improving targeted marketing campaign effectiveness by **18%**.
* Migrated ML workloads from **on-prem to GCP Vertex AI**, cutting operational costs by **25%**.

**Data Analyst | Fortis Healthcare | May 2017 – Jun 2019**

**Industry: Healthcare | Predictive Analytics**

* Built predictive models to identify **high-risk patients for chronic diseases** using logistic regression and random forests, enhancing early diagnosis rates by **30%**.
* Developed a **patient readmission prediction system** for hospitals, reducing readmission rates by **12%**.
* Integrated data pipelines from **electronic health records (EHR)** into a unified data lake, improving data accessibility for analytics teams.

**Key Projects**

**1. Real-Time Fraud Detection for Fintech Payments**

* **Tech Stack:** Kafka, Spark Streaming, Python, Scikit-learn, AWS Lambda, MongoDB
* Developed a **streaming fraud detection system** that **processes 1M+ transactions/day** and flags suspicious activities in real-time.
* Used **unsupervised learning (Autoencoders, Isolation Forests)** to detect anomalies and reduce fraud losses by **$5M/year**.

**2. Predictive Maintenance for Manufacturing Industry**

* **Tech Stack:** PySpark, TensorFlow, PostgreSQL, Power BI
* Built an **AI-driven predictive maintenance model** for a large manufacturing plant, reducing downtime by **25%** and maintenance costs by **$2M annually**.
* Implemented an **IoT sensor data ingestion pipeline** for real-time failure prediction.

**3. E-commerce Customer Churn Prediction**

* **Tech Stack:** Python, XGBoost, Flask, AWS
* Developed a **customer churn prediction model** using XGBoost, achieving an **86% accuracy** and enabling targeted retention strategies.
* Integrated with a **Flask API for real-time prediction and dashboard visualization**.

**Education**

**Bachelor’s in Computer Science**  
Indian Institute of Technology (IIT) Bombay, 2017

**Certifications**

* **AWS Certified Machine Learning – Specialty**
* **Google Professional Data Engineer**
* **Microsoft Certified: Azure AI Engineer Associate**
* **IBM Data Science Professional Certificate**

**Achievements & Publications**

* Published a research paper on **“Fraud Detection in Fintech Using AI”** in IEEE Transactions.
* Speaker at **PyData Conference 2023**, presenting work on **Real-time ML in Fintech**.
* Developed an **open-source library for feature engineering in Python** with **1K+ GitHub stars**.

**Personal Projects & Open Source Contributions**

* Developed an **AI-powered chatbot for customer service automation** using **RAG-based Gen AI**.
* Contributed to **Apache Spark MLlib**, improving certain algorithms for large-scale processing.
* Built a **Python package for automated ML model evaluation** with an intuitive dashboard.

**Soft Skills**

* Strong problem-solving & analytical skills
* Effective communication & storytelling with data
* Agile & Scrum methodologies experience
* Leadership & mentoring junior data scientists