



Pranjal Shukla

Mob: (+91) 9910631251 | pranjalshukla98@gmail.com
linkedin/pranjalshukla98 | github.com/pranjals26



EDUCATION

Course	College/University	Year
Master's of Business Analytics - AI & Cloud	George Washington University	Aug 2022 - Dec 2023
B.Tech in Computer Science	Vellore Institute of Technology	June 2016 - July 2020

CERTIFICATIONS

- **AWS Certified:** Cloud Practitioner & Solution Architect
- **DataBricks Certified:** Data Engineer

WORK EXPERIENCE

- **Data Science Manager - Digital Innovation | Shriram Finance | Bangalore, India** [Dec. 2024 - Present]
 - Spearheaded a **7-member team**, providing technical mentorship and project oversight to ensure high-quality delivery within the AI & Data Science division.
 - **Agentic RAG SME:** Architected a multi-tool Fixed Deposit (FD) Agent using AWS Bedrock and Strands SDK with function calling for real-time financial data retrieval via Crawl4AI. Built a semantic routing framework to resolve complex user queries while ensuring regulatory compliance and data freshness. **Delivered a 15% uplift in FD conversions (298Cr/Month).**
 - **Serverless FAQ Bot:** Deployed a highly scalable RAG agent using AWS Bedrock, Lambda, and API Gateway. Implemented a hybrid retrieval logic that routes static queries to SQL databases and complex queries to a Vector Knowledge Base, resulting in a **10% increase in lead generation** for leading & investment products.
 - **LLM Fine-Tuning:** Architected an end-to-end SMS Engine using Gemma 3 27B fine-tuned with Unsloth. Developed a multi-stage pipeline from noise removal to cohort generation, processing **60M+ raw SMS** to identify **50+ financial and behavioral segments**. Automated daily ingestion to on-prem SQL servers to directly power MoEngage and Netcore marketing campaigns.
 - **Predictive Modeling & Collection ROI:** Reduced CSPL EMI bounce impact by **capturing 80% of high-risk accounts** with a CatBoost-based bounce prediction classification model. Engineered **70+ features** enabling prioritized collection and **driving 70 Cr in incremental monthly arrears.**
- **Data Solution Architect | Latent Plane | London, UK** [Sep. 2023 - Dec. 2024]
 - **Resilient Async Infrastructure:** Engineered an asynchronous event framework using AWS Lambda to handle high concurrency, **guaranteed 95% query resolution**, and **<2s latency**.
 - **GenAI Cost Engineering:** Architected a GenAI agent projected to **cut annual support costs by ~75%**, replacing a **\$55k-\$75k** human-staff model with a **\$12k-\$15k** automated system, while also slashing operational costs by designing a hybrid RAG strategy to reduce expensive LLM API calls.
 - Developed an intelligent orchestration engine utilizing Claude 3 for real-time intent and sentiment analysis, automating routing and escalation via MongoDB/SES while maintaining a **6-year compliance lifecycle** through automated S3 data archiving.
- **Product Manager Intern | Datalink Software | Tampa, Florida, USA** [May 2023 - Sep. 2023]
 - **Data Pipeline Engineering:** Implemented a Databricks ETL pipeline to process and standardize **10,000+ patient records**, significantly enhancing regulatory compliance and medication adherence reporting.
 - **Technical Delivery Management:** Led the end-to-end delivery of **12 client features** by translating **50+ business use cases** into technical specifications, **improving project execution efficiency by 25%**.
- **Technical Associate | Intellex Investment Solutions | Noida, India** [June 2020 - Aug. 2022]
 - **Cloud Migration & Automated ETL:** Engineered a streamlined data workflow by migrating on-prem legacy data to AWS RDS (DMS). Configured serverless ETL pipelines using AWS Glue, Kinesis, and Step Functions, **reducing data entry errors by 10%**.
 - **Financial Risk Modeling:** Built a credit risk classification model for a **30,000-customer portfolio**, achieving **85% accuracy** in predicting payment delinquency to inform credit line strategies.

SKILLS & INTERESTS

- **AI Machine Learning :** Machine Learning, Deep Learning, Natural Language Processing (NLP), Prompt Engineering, Transformers, Classification Models (XGBoost, LightGBM, CatBoost, Naive Bayes), PyTorch, Keras, scikit-learn
- **Generative AI, LLM Frameworks Tools:** LangChain, RAG-based architectures, Prompt Engineering, AWS Bedrock, Agent Core, Strands, OpenSearch, Pinecone, FAISS, AWS SageMaker Model Deployment Strategies (A/B Testing, Blue/Green, In-Place), Claude Code, Claude Flow, Microservices Architecture, Python, FastAPI, REST APIs, SQL, DataPipeline
- **Product Strategy :** Product Management, Business Requirements Analysis, Strategy Development, Product Analytics
- **Leadership Stakeholder Management :** Team Management, Stakeholder Management, Cross-functional Collaboration