

Dushyant Singh – Risk & Decision Scientist | AI Engineer

Pune, India | +91 7014093420 | dushyant.singh.civ16@iitbhu.ac.in

LinkedIn: [linkedin.com/in/dushyant-singh-3214a8144](https://www.linkedin.com/in/dushyant-singh-3214a8144) | GitHub: github.com/dushyant4342

Professional Profile

I am a results-driven AI and data professional with more than four years of experience converting complex data into business impact across risk analytics, data engineering and generative-AI product lines. My work spans full-stack AI delivery—from architecting AWS-scale pipelines and MLOps governance to researching open-source LLMs and mentoring teams through deep-learning bootcamps. A fast learner and hands-on leader, I thrive on shipping production systems, communicating clearly with stakeholders and staying on the cutting edge of models such as DIA, Llama 3, DeepSeek, Mistral and D-Fine.

Experience

AI Engineer – FPL Technologies (OneCard) — Feb 2025 – Present

Since early 2025 I have spearheaded OneCard's GenAI initiatives. I engineered a multilingual intent-classification service for our customer chatbot, pairing HuggingFace transformers with Qdrant semantic search to slash external API costs by 70 percent. In parallel I delivered a production-grade voice bot that chains Whisper ASR, LangChain routing and Coqui TTS; WebRTC-VAD keeps the conversation latency-free while our telephony stack handles thousands of daily calls. To scale our talent, I ran a six-week Deep-Learning & GenAI bootcamp for fifty-plus engineers, walking them through ANNs, CNNs, LoRA fine-tuning, MCP servers and agentic RAG design. I also publish weekly Reddit-sentiment dashboards benchmarking OneCard against competitors and built an NLP pipeline that mines customer-call remarks to fine-tune OneScore credit models and drive better future-call recommendations.

Lead – Data Management – FPL Technologies (OneCard & OneScore) Mar 2024 – Jan 2025

During this period I reorganised the firm's data estate into L1/L2/L3 Delta Lake zones, migrated terabytes with PySpark-driven AWS Glue jobs and wired Starburst Trino to unified S3 views. GitLab CI/CD on EC2, combined with Lambda-orchestrated cron deployments, gave ninety data users a seamless—and cheaper—path to governed data. I launched an AWS Feature Store with MLflow drift monitoring (KSI/PSI) surfaced in Superset and alerted via Google Chat. Superset itself was migrated to Trino on EC2, upgraded for PDF exports and linked to OpenMetadata for lineage, tests and profiling. I further secured PII with Ranger masking, automated SFTP/API/Jupyter workflows through Lambda and stood up Kinesis-Firehose-EventBridge streams that pre-process callback payloads in flight.

Risk & Decision Scientist – FPL Technologies (OneCard) — Nov 2020 – Aug 2023

Earlier, my focus was risk modelling and collections strategy. I built an XGBoost segmentation model (81 % Gini) that lifted pre-due-date resolutions from 89 % to 92.4 % across a million Bucket-0 accounts, designed triggers and campaigns that drove early-delinquency rates down from 3.18 % to 2.54 % on a ₹3 000 Cr portfolio, and created propensity-to-pay and agency-allocation models that improved recovery by double-digit percentages while curbing NPA flow. Random-Forest channel optimisation raised NPS by seven points, and the OneAccord in-house collection platform I co-product-managed now unifies dialling, nudges and billing under a secure, auditable interface.

Skills

Advanced Python, SQL and PySpark underpin my model-driven analytics, while XGBoost, Random Forests, LightGBM and deep nets (PyTorch/TensorFlow, ANN/CNN/RNN) power both classical risk work and transformer-based LLM experiments (Llama 3, Gemma, Qwen, DeepSeek, Phi-3). I build conversational and vision systems with Whisper, Coqui TTS, YOLOv8 and D-Fine, orchestrate RAG workflows via LangChain + LlamaIndex over Qdrant, and productionise on FastAPI, Docker and AWS with MLflow, CI/CD, OpenMetadata lineage and Ranger security. My soft-skill toolkit spans leadership, cross-functional storytelling and hands-on deep-learning workshops.

Key Projects & Deployments

I have containerised and published a Bedrock-powered document-QA app (Titan Embed + Llama 3) at universal-app.info, shipped a DeepSeek R1 1.5 B chat application via Ollama and Docker Hub, and previously fine-tuned VGG16 and Fashion-MNIST classifiers to ~90 %+ accuracy. Earlier RNN-LSTM fake-news and ANN churn models rounded out a solid foundation in NLP and predictive analytics.

Achievements & Awards

The **OneAccord Award** recognised my two-month delivery of the in-house collection system, and IIT-BHU's Gymkhana Award cited my contribution to the institute's Film & Media Council.

Education & Interests

B.Tech., Civil Engineering — IIT BHU Varanasi (2016 – 2020)

Outside work, I enjoy music, drone-cinematography, photography, swimming and hiking, pursuits that keep my curiosity—and batteries—fully charged.