

MANISH SHUKLA

📍 Dallas, TX ✉ manishshukla.ms18@gmail.com ☎ 4693803183 🌐 in/manishshukla-ms 🌐 www.manishshukla.com

EXPERIENCE

Lead Gen AI Data Scientist | SAVVYAN TECHNOLOGIES, VERIZON | US, TX, Irving | January 2024 – Present

- Developed a Large Language Model (LLM) solution using PaLM and a tree-based prompting algorithm to classify 75K+ documents into clause categories—saving \$2M annually and eliminating the need for 4 FTEs by replacing legacy systems with an in-house solution.
- Formulated contract document classification model with interpretability via LIME/SHAP, conducted class-level evaluation and proposed roadmap with BERT-based improvements, human review, and production deployment planning.
- Built a proof of concept using the multi-agent LangGraph framework (CFO, Supervisor, Critique, and Writer agents) to analyze customer connectivity service usage end-to-end and generate actionable recommendations—projected to deliver \$4.6M in annual savings.
- Designed and deployed an automated Newsletter AI using LangGraph and MCP tools, orchestrating News Gatherer, Critique, and Chief Editor agents to streamline content creation and distribution to 1,000+ stakeholders.
- Created a framework leveraging LLMs to generate synthetic data for secure data handling and model evaluation, supporting the benchmarking of various LLMs and third-party solutions for enterprise use cases.
- Led a team of 5+ data scientists and analysts, driving project execution, aligning on deliverables, and defining development workflows to ensure timely and high-quality outcomes.

Senior Data Scientist | DELOITTE | US, TX, Dallas | January 2022 – November 2023

- Developed an end-to-end time series forecasting system for user-related KPI metrics across wallet payment platforms in multiple markets, enabling data-driven strategy and planning decisions to boost product usage and revenue.
- Built a robust anomaly detection system for KPI metrics reported by 2K+ partners, improving reporting timeliness and reducing 80+ hours of manual effort each month.
- Led cross-functional teams of Data Scientists, UI Developers, and Data Engineers to automate the MLOps pipeline, including data ingestion, cleaning, model development, visualization, and deployment via cloud platforms and Hadoop (Scala-Spark).
- Designed a proof of concept for text summarization of 1,000+ retail product files using Generative AI techniques (PaLM, prompt engineering, NLP/NLU), delivering insights to enhance customer sentiment analysis and improve purchase behavior.
- Implemented an intelligent chatbot for a retail client to handle queries for 500+ products, using advanced NLP/NLU techniques and fine-tuned LLMs such as Falcon-40B, Falcon-7B, LLaMA 2, and BERT.

Data Scientist & Technical Advisor | PALLADIUM | US, TX, Dallas | October 2020 – January 2022

- Developed a Time Series model to forecast the number of positive HIV & AIDS cases in African countries at Local Municipality level using exogenous factors to benefit 30+ millions patients.
- Performed Media Analytics to find out effectiveness of training conducted by USAID for Women Empowerment and Domestic Violence Awareness.
- Led the team of 10+ Data Scientists and Analysts to guide on the deliverables, meet the goals and define the process for development.

Data Scientist | TECHSTAR, VERIZON | US, TX, Irving | December 2019 – May 2020

- Built an AI/ML model using classification algorithms to predict which server to be patched for 5G up-gradation, benefited \$1.7 million.
- Interpreted prominent features by weight of evidence from 200+ features for more than 28 million services to provide actionable insights for services mismatch and failure in up-gradation.

Data Science Intern | TRUEVIM | US, TX, Lewisville | August 2019 – November 2019

- Developed a Time series ARIMAX and VAR model, which helped to forecast house prices by considering exogenous variables such as Dow Jones, Unemployment Rate, Interest Rate and Consumer Sentiment Analysis.
- Built regression models using Random Forest and Neural Networks for extracted 5+ million raw data by using Beautiful Soup to predict house price and deployed using Flask to create proof of concept API.

Data Science and Analyst | L&T Infotech | IN, MAHARASHTRA, Mumbai | June 2015 – June 2018

- Added 450+ functionalities, 100 suggestions to improve stability and automated daily tasks, which resulted in savings of \$ 12 million, system stability up to 60% and averted 20+ hours of manual efforts on weekly basis.
- Built a classification model using Python to anticipate the buy probability of a car model for customer, and interpreted the features that highly influence the purchase of a car, which resulted in 16% profit.
- Implemented a Logistic Regression to find that whether consumer bought the vehicle again or not and found factors dealing with selling the car most, which improved sales by 9%.

EDUCATION

M.S., Business Analytics (Data Science) | The University of Texas at Dallas | Dallas | 2020

B.E., Electronics & Telecommunication Engineering | University of Mumbai | Mumbai | 2015

SKILLS

- Machine Learning (Supervised/Unsupervised) • Time Series Forecasting (ARIMA/VAR) • Languages : Python, PySpark, Scala, R, SAS, Pytorch
- Database Systems (SQL/NoSQL) • Statistical and Exploratory Data Analysis • Github/Git, Airflow • Google Analytics / Adobe Analytics
- Generative AI and Large Language Models such as PaLM, BERT, etc • Anomaly Detection & Business Intelligence • Tableau/PowerBI

CERTIFICATIONS

IBM Data Science Professional | GCP Professional ML Engineer | GCP Generative AI | GCP Digital Cloud Leader