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 LLM system using PALM and tree-based prompt algorithm to classify 75k+ document in clauses category, which saved $2M on yearly basis and 4 FTEs manual efforts by replacing traditional system and providing in house solution.
· Built a proof of concept using the Multi-agent LangGraph framework, which includes CFO, Supervisor, Critique and Writer agents, to perform end-to-end analysis for the customer connectivity service usage and provide recommendations, which could save $4.6M annually.
· Created a framework that utilizes LLM models to generate synthetic data for protecting and managing original data, which aided in evaluating various LLM models and third-party services for resolving business use-cases.
· Led a team of more than 5 data scientists and analysts to ensure the deliverables were met, goals were reached, and the development process was defined.

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

· Developed an End-to-End Time Series system to forecast User related KPI metrics for Wallet payment, strategy and planning purpose in different markets, which helped to suggest actionable insights for improving the product usage and revenue.
· Built a complete system to detect anomaly in User related KPI metrics reported by 2K+ partners for planning purpose, which led to report metrics on time and reduced 80+ manual hours monthly.
· Followed Software Development Life Cycle process for requirement gathering and sign off, Design the solutions, Development, User Acceptance Testing(UAT), Deployment and Scalability using Pytorch.
· Led teams of Data Scientists, UI Developers and Data Engineers to automate the ML Ops process from data collection and cleaning using CI/CD Pipeline, model building, visualization to deployment using Cloud platforms and Hadoop jobs with Scala-Spark.
· Designed a POC for Text summarization from 1000+ files to analyze sentiment on retail products using Generative AI prompt engineering concepts, NLP/NLU techniques and PaLM LLM modeling for improving the sales and purchase power.
· Implemented a chatbot for retail client to answer the queries related to 500+ products using NLP/NLU techniques, prompt engineering and fine tuning LLM model 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 in Society.
· 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