MANISH SHUKLA

Senior Data Scientist

+1-469-380-3183 · manishshukla.ms18@gmail.com · Dallas, TX

Skills

- Machine Learning (Supervised/Unsupervised)
- Time Series Forecasting (ARIMA/VAR)
- Languages : Python, PySpark, Scala, R
- Tableau/PowerBI
- Database Systems (SQL/NoSQL)
- Statistical and Exploratory Data Analysis
- Anomaly Detection and Predictive Modeling
- Generative AI
- Google Analytics / Adobe Analytics
- Github/Git, Airflow

Experience

Senior Data Scientist
DELOITTE, Dallas, TX
January 2022 - Present

- 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
- 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 concepts and LLM modeling for improving the sales and purchase power

Data Scientist & Technical Advisor
PALLADIUM, Dallas, TX
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 Data Scientists and Data Analysts to guide them on the deliverables, meet the project goals and define the process for development

Analytics Intern
MARY KAY INC, Dallas, TX
June 2020 - October 2020

- Implemented a prediction model to find the probability that user will click on email advertisement based on the features of the user and establish metrics to drive the product campaigns

Data Scientist
TECHSTAR, CLIENT : VERIZON, Irving, TX
December 2019 - May 2020

- Built an AI/ML model using classification algorithms and web-scraped data to predict which server needs to be patched for 5G upgradation, resulting in saving of $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 upgradation

Data Science Intern
TRUEVIM, Lewisville, TX
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

Analytics Intern
MARY KAY INC, Dallas, TX
May 2019 - August 2019

- Performed Segmentation on 2+ Million customers with K-means clustering to target customers in the United States and redefined marketing strategies, which increased conversion rate by 9%
- Designed a POC for a chatbot using Natural Language Processing and Flask to shorten the communication gap between beauty consultants and consumers about the products details

Data Science and Analyst
LARSEN & TOUBRO INFOTECH LTD, LTI, CLIENT : HONDA NORTH AMERICA, 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 will buy the vehicle again or not and found factors dealing with selling the car most, which improved sales by 9%

Education

M.S in Business Analytics - Data Science
The University of Texas at Dallas
May 2020

B.E in Electronics & Telecommunication
University of Mumbai, US
June 2015

Certifications

- IBM Data Science Professional
- GCP Generative AI
- GCP Professional Machine Learning Engineer
- GCP Digital Cloud Leader

Awards

- Mary Kay WOW Awards
- Valuable Performance Award
- Forbes Under 30 Scholar
- MLH Hackathon Winner

Linkedin

https://www.linkedin.com/in/manishshukla-ms/

Website

https://www.manishshukla.com

Github

https://github.com/Manishms18