First Last Name

✉ email@gmail.com | 📱 +44-phone | in linkedin-gk

SUMMARY

Data Scientist with expertise in data analysis, machine learning, predictive modeling, and software engineering. Able to effectively communicate technical concepts to non-technical stakeholders and present insightful data-driven solutions to drive business growth. Proficient in Python, SQL, and statistical analysis. Specializes in NLP with a focus on document classification.

EDUCATION

MSC DATA SCIENCE, UNIVERSITY OF SUSSEX (2021-22)
B.TECH (IT), AMITY UNIVERSITY(2012-16)

EXPERIENCE

JAN 2022 – JUN 2022
Research Assistant, **SPRU (University of Sussex)**

- Designed architecture to ingest about ~700 GB citation data from Open Alex and perform analysis using Azure Databricks, resulting in a 20% increase in data processing efficiency.
- Led a team of 4, implementing Leiden clustering algorithm for citation clustering using Python and Java with Sussex HPC.

JAN 2018 – AUG 2021
DATA SCIENTIST/SENIOR SOFTWARE ENGINEER, **Coforge Ltd.**

- Automated manual data extraction process of classifying vehicle insurance documents from multiple dealerships, improving SLA from 50 minutes to 1 minute.
- Orchestrated a speech analytics system using Apache NiFi, converting lingual audio to text, transcribing to English, and performing sentiment analysis using NLP and Deep Learning.
- Collaborated with data analysts and team leads to create solutions increasing product growth by 18%.
- Designed forecasting algorithms to predict the maximum potential revenue for an Ad-tech performing real-time auctions and setting the reserved price for these auctions, leading to a 15% increase in auction revenue.
- Streamed ATM transaction data using Kafka and Databricks and generated outcomes for a reporting dashboard, providing stakeholders with real-time visibility into transaction data.
- Deployed a text classification model using Flask to assign IT support tickets to their respective groups, achieving an accuracy of 81%, and f-score of 0.8, and reducing response time by 50%
- Conducted exploratory data analysis for Customs Risk Engine for Dubai Customs (DUTECH) and was extensively involved in writing SQL queries, resulting in a 20% improvement in data quality.

OCT 2016 – JAN 2018
SOFTWARE ENGINEER, **NIIT Technologies**

- Conducted A/B testing, for e-commerce platform to understand the impact of user behavior, increased coupon code utilization by 15%.
- Performed data migration from on-premises SQL Server to Amazon EC2 instance.
- Developed a Ticket Reservation System for the SITA (Airline Career), ensuring high-quality code in an Agile environment.

TECHNICAL SKILLS

- Programming: Python, REST APIs, SQL
- Machine Learning Libraries: scikit-learn, TensorFlow, Keras
- Data Visualization and Analysis Tools: Jupyter Notebook, Power BI
- Cloud: Azure, Azure Databricks, Azure Synapse, Azure Datalake

ADDITIONAL CERTIFICATIONS

- AWS Certified Developer-Associate
- Data Science with Python (AnalytixLabs)

OTHER WORK

- Masters' dissertation thesis on Long Document Classification comparing hierarchical transformer-based models.
- Conference Paper on Solving Traffic Congestion – An application of VANET.