SUMIT GUNDAWAR

+44 7810 081335 | sumitgundawar3@gmail.com | London, England, United Kingdom |

linkedin.com/in/sumit-gundawar-759470129/ | github.com/sumitgundawar | www.sumitgundawar.com/

PROFESSIONAL SUMMARY

Data Scientist and Analyst with 3+ years of experience in business analytics, statistical analysis, and data-driven decision-making. Proven track record in delivering actionable insights through advanced analytics and visualization. Strong background in developing and implementing data solutions using Python, SQL, and cloud technologies. Demonstrated expertise in translating complex business problems into analytical solutions while effectively communicating findings to stakeholders.

EDUCATION

University of East London   **September 2023 - September 2024**
Master's, Data Science   *GPA: 81*

Vellore Institute of Technology   **July 2019 - June 2021**
Master's, Computer Science   *GPA: 9.26*

Savitribai Phule Pune University   **June 2016 - June 2019**
Bachelor's, Computer Science   *GPA: 87.03*

CERTIFICATIONS

• Applied AI with DeepLearning (Coursera – Jan 2024)
• Advanced Machine Learning and Signal Processing (Coursera – Sept 2023)
• Analytics Bootcamp Program (Great Learning – Jul 2021)
• Machine Learning with Python (Foxmula – Jan 2021)
• Introduction to TensorFlow for Artificial Intelligence, Machine Learning, and Deep Learning (Coursera –Oct 2020)
• AWS Cloud Practitioner Essentials (AWS – Jul 2020)
• Google Cloud Platform Fundamentals Coursera – May 2020)
• AI From the Data Center to the Edge (Intel – Mar 2020)

PROFESSIONAL EXPERIENCE

LatentView Analytics   **Chennai, TN, India**
Data Analyst   *June 2021 - June 2023*
- Architected and deployed production-grade machine learning systems, focusing on scalability and high-throughput prediction capabilities
- Developed robust ETL pipelines using PySpark for processing large-scale data across multiple cloud environments
- Implemented infrastructure-as-code using Terraform for ML model deployment across Azure and GCP
- Built monitoring systems for ML model performance and business impact metrics using PowerBI and Tableau
- Mentored junior team members and collaborated with cross-functional teams to drive ML initiatives

KEY PROJECTS

Picker Accuracy   **Chennai, TN, India**
Data Analyst   *June 2021 - September 2021*
- Developed a custom algorithm in Python to calculate and improve picker accuracy in PepsiCo's North
- American warehouses, replacing a legacy system and improving efficiency by 70%.
- Leveraged Azure Data Lake and Presto to optimize the process of tracking and analyzing picker performance, reducing the time required for metrics calculation.
- Implemented a causal inference approach to identify factors affecting warehouse picker performance, enabling PepsiCo to reduce human error and optimize employee workflows.
- Improved the overall warehouse operations by providing actionable insights, contributing to enhanced operational efficiency across the logistics network.

Demand Planning Disaggregation   **Chennai, TN, India**
Data Analyst   *September 2021 - June 2023*

- Designed and developed an 8-week demand forecasting model using SAS and PySpark, achieving a 72% accuracy rate for forecasting demand across 25 million customer-product combinations in North America.
- Built and managed ETL pipelines on Azure Databricks for real-time data ingestion and processing, optimizing the storage of transformed data in Azure Data Lake.
- Integrated statistical techniques such as regression analysis and time-series analysis to improve predictive accuracy, enhancing overall supply chain forecasting for Frito Lay products.
- Generated actionable insights through Power BI dashboards, providing executive leadership with data-driven decisions that contributed to a $1.2B annual revenue impact.
- Led cross-functional collaboration with teams across regions to fine-tune the model and align it with PepsiCo's global business objectives.

Chicken Disease Classifier **London, England, United Kingdom**
January 2024 - February 2024
- Developed a deep learning-based chicken disease classification model, with streamlined workflows using DVC for data version control and pipeline reproducibility.
- Configured and optimized CI/CD pipelines with GitHub Actions, automating the build, test, and deployment processes for both AWS and Azure environments.
- Deployed Dockerized application on AWS using EC2 for compute and ECR for container storage, ensuring efficient management and scalable access to the classification service.
- Set up a secondary deployment pipeline on Azure, utilizing Azure Container Registry and Web App services to securely manage and deploy Docker images for web-based access.
- *Link to project*

Low Velocity Item Ordering **Chennai, TN, India**
Data Analyst *November 2021 - January 2022*
- Conducted advanced data analysis using Python and PySpark, identifying optimization opportunities for slow-moving products in retail stores.
- Utilized historical sales and inventory data to build predictive models that helped streamline inventory management, resulting in an improvement in product turnover rates.
- Collaborated with product managers to deploy optimization algorithms that improved order planning, leading to a $24M annual revenue increase by reducing stockouts and over-ordering.
- Automated the generation of Power BI reports, enabling the supply chain team to continuously monitor product performance and quickly adjust strategies based on real-time data.

Stock Price Tracker **London, England, United Kingdom**
August 2023 - September 2023
- Developed a real-time stock price tracking application using Flask and JavaScript, allowing users to monitor current prices and percentage changes for stocks, index funds, ETFs, and cryptocurrencies.
- Integrated the Yahoo Finance API to retrieve live financial data and implemented a 10-second update interval to ensure real-time data visualization on the web interface.
- Built a recommendation feature using a CSV file lookup, enhancing user experience by providing symbol suggestions during asset search.
- Designed a color-coded display to visually indicate asset price changes (green for positive, red for negative, gray for no change), improving data clarity and user engagement.
- *Link to project*

SKILLS

Programming Languages & Tools: Python, SQL, R, Java, NoSQL, MySQL, Postgres
Data Science & Machine Learning: Data Science, Data Analysis, Natural Language Processing (NLP), Deep Learning, Pytorch, Tensorflow, Keras, Python NLTK, SpaCy, Mathematics
Big Data & Data Processing: Apache Spark, Apache Kafka, Apache Hive, BigQuery, Databricks, Hadoop, Google Cloud Platform, Microsoft Azure, Data Lakes, ETL Pipeline Development, Real-time Processing
Cloud & DevOps: Docker, Kubernetes, Terraform, Airflow, Jenkins, Data Warehouse
Analytics & Visualization: Power BI, Tableau, Amazon Quicksight, Google Analytics, Statistical Analysis
Data Manipulation & Libraries: NumPy, Pandas, RDBMS, MongoDB, SQL, Polars

WORK AUTHORIZATION

Eligible to work in United Kingdom on a Post-Study Work (PSW) visa. Open to sponsorship for long-term roles.