

Abhishek Gupta

abhishek.gupta0724@gmail.com | <https://abhishek-xavi.github.io/> | 07833460266 | England

PROFILE

Machine Learning Enthusiast with over 4 years of experience in software delivery process, data consultant and cloud infrastructure in Retail domain. Innovated advanced analytics solutions to enhance operational functions leading to 20% reduction in service incidents and savings of \$500K per year. A data scientist with analytical, statistical, and programming skills looking to develop data-driven solutions and increase business revenue.

EDUCATION

Master's in Data Science

10/2022 – 09/2023

Lancaster University, England

- Developed facial image to sketch model, leveraging Convolutional Autoencoder and generative adversarial network (GAN) with attention mechanisms in a U-net neural network. Data transformation was leveraged for experimentation achieving a 72% reconstruction accuracy with GAN.
- Designed ETL tasks to collect raw data for data analytics and machine learning process from real time google APIs. A relation database management system was designed and setup on Hadoop with a data visualization dashboard to identify data patterns and trends using Apache NiFi and Grafana.
- Researched on Lancashire Crime data in collaboration with GBG, employing advanced analysis and time series forecasting to identify crime hotspots and provided data-driven solutions.

WORK EXPERIENCE

Data Scientist

06/2023 – 09/2023

Experian, London

- Performed NLP analysis on over 50,000 Notice of Correction documents to identify financially responsible consumers to prevent fraud and optimize credit risk assessment.
- Responsible for project facilitation, data transformation, and the development of end-to-end Machine learning project in python. Classification models including Logistic Regression, Random Forests and XGBoost were experimented and analyzed.
- Planned the tasks with accuracy and broke down large tasks into smaller incremental deliverables that can be completed in a single sprint. This involved data collection, preprocessing, feature engineering, and model development.
- Implemented self-supervised learning models such as LDA and NMF algorithms to extract meaningful insights and decompose consumer habits and behaviors to fulfill business requirements.
- Collaborated with the Data Science and Machine Learning engineers team and achieved 72% classification accuracy in identifying financially responsible consumers.
- Improved precision of existing credit risk model by 8% through integrating enhanced NLP-based consumer profiling.

Data Science Intern

01/2022 – 07/2022

Crux Intelligence, India

- Developed conversational chatbot with Dialogue Management using Bot SDK framework in Python and integrated with Microsoft teams. This led to 30% increase in the APIs usage.

- Utilized ANTLR (Another Tool for Language Recognition) to develop code for efficient analysis of business KPI's, enhancing feature relationship comprehension through code-generated regularization.
- Optimized Natural Language Processing module performance, using Python parallelization libraries like Gunicorn. Refactored module codebase making it cost effective by 80%, and performance optimized for cloud infrastructure on Kubernetes.

Technical Analyst

06/2017 – 07/2021

Blue Yonder, India

- Demonstrated effective leadership and collaborated closely with business and IT stakeholders. The responsibility involved designing technical solution and ensuring it fulfills business needs.
- Developed an analytics platform to automate monitoring of 100+ ETL pipelines, empowering advanced analytics and identifying data failures. This Reduced service incidents by 20% resulting in \$500K savings.
- Spearheaded deployment of category management solution for \$30B revenue customer. The responsibility involved collaboration with cross-functional teams, planning, testing, implementing, and deploying, while interacting with external and internal stakeholders.
- Led root cause analysis for top retailers, big data issues utilizing SQL profiling. Reduced incidents by 30% through query plan executions and optimizing the code.
- Responsible for data engineering and automated ETL data pipelines through the development and optimization of Batch and PowerShell scripts, ensuring data quality and data security.
- Exhibited strong presentation abilities through delivering 10+ training sessions on proprietary solutions to overseas partners in China and India. Clearly conveyed technical knowledge to non-technical audiences.
- Mentored internal team members on software engineering principles, cloud infrastructure, corporate protocols, and client meetings.

SKILLS

Technical Skills: Python | R | SQL | Oracle | Scripting Languages | Linux | Algorithms | Data Structure

ML Concepts: Regression | Classification | Clustering | Ensemble methods | Word embedding | Natural Language Processing (NLP) | Reinforcement Learning (RL) | Neural Network | Deep Learning | Machine Learning | Statistical Analysis| Predictive models | Artificial Intelligence

Libraries: Pandas | Numpy | Tensorflow | Keras | Seaborn | sci-kit learn | NLTK | Transformer | sPacy

Tools: GitHub | git bucket | Jira | salesforce | AWS | Docker | Kubernetes (K8) | Azure