Ajinkya Chandrakant Bhuse

- M. Sc. in Data Science
Aspiring Data Scientist
Guildford, UK | +44-7909304583
ajinkyacbhuse@gmail.com| www.linkedin.com/in/iamajinyabhuse/ |www.github.com/ajinkyabhuse3798

PROFILE SUMMARY

- Results-driven individual pursuing **Master's in Data Science**, equipped with **2.3 years of experience** in system engineering, seeks to transition into a dynamic role as a **Data Scientist.**
- Keen analytical mindset and a passion for extracting meaningful insights from complex datasets.
- Experienced in developing complex data models using diverse datasets, including structured, semi-structured, and unstructured data.
- Adept at analyzing data to identify patterns, trends, and anomalies, and assess suitability for various ML techniques.
- Proven ability to develop & execute robust plans for data analysis, outlining methodologies & evaluation criteria.
- Skilled in creating and maintaining insightful reports for internal and client-facing audiences, effectively communicating data-driven recommendations and informing strategic decision-making.
- Capable of driving the e2e process of gathering, analysing, and delivering business data to establish KPIs.
- Experienced in utilizing machine learning algorithms for predictive modelling and resolving real-world challenges.
- Strong communication skills and a collaborative approach to problem-solving, coupled with a commitment to continuous learning and professional growth.

EDUCATION

MSc Data Science **2023**
University of Surrey, Guildford, United Kingdom

Modules:
Data Science Principles and Practices || Computational Intelligence || Practical Business Analytics || DBMS || Cloud Computing || ML and Data Mining || Natural Language Processing || Advanced Challenges in Web Technologies

BE in Electronics and Telecommunication **2020**
SVERI's College of Engineering, Pandharpur
Received Tuition Fee Waiver Scheme for College and University Students MHRC, India

EXPERIENCE

Senior System Engineer | Data Analyst
Infosys Ltd
May 2021 – Aug. 2023

- **Complex Data Analysis:** Contributed to complex data analysis initiatives aimed at identifying various opportunities and root causes, resulting in a reduction of customer effort and an improvement of metrics by approximately **25%.**
  → Analysed **10GB** of data, pinpointing growth opportunities & implementing record changes that led to an annual saving of **$500k.**
- **KPI Generation:** Generated, forecasted, and established Key Performance Indicators (KPIs), providing strategic insights pivotal for driving growth and enhancing performance.
- **SQL Query Development:** Created over **3000 SQL queries** to fulfil diverse technical data requirements for projects and business reviews, ensuring comprehensive data access and analysis.
- **Dynamic Dashboard Development:** Developed dynamic dashboards based on KPIs using Tableau, enabling clients to monitor project status in real-time.
- **Cross-functional collaboration:** Collaborated with cross-functional teams in formulating strategies, facilitating interactions, and coordinating with onshore counterparts to meet project requirements effectively.

SKILLS

Functional Skills: Data Modelling, Big Data Technologies, Predictive Analysis, Machine Learning, Data Governance, Deep Learning, Statistical Analysis, Root Cause Analysis, Stakeholder Collaboration

Technical Skills: Python, AWS, Tableau Google Cloud, SQL, CI/CD Pipelines, AWS Sagemaker, Docker, GitHub, MS Excel

Algorithms: Linear Regression, Decision Trees, Logistic Regression, K-Means, KNN, Hierarchical Clustering, SVM, Ensembles, Tree-Based Models, Naive Bayes, Neural Networks, Transformers

PROJECTS

Optimization of Neural Network for CIFAR-10 Image Classification | Oct. 2023 – Dec. 2023

Tools Used: Python, Convolutional Neural Network (CNN), Genetic Algorithm, Gradient Descent, Differential Evolution

- Developed a solution for classifying images from the CIFAR-10 dataset, enhancing the accuracy of image classification using neural networks.
- Selected an appropriate neural network architecture tailored for image classification tasks, implementing the network in Python.
- Implemented a two-phase training approach, initially employing a standard gradient descent algorithm to establish a baseline performance for the neural network.
- Introduced a novel approach by randomizing the weights of the last layer and initiating training from scratch.
- Applied advanced optimization algorithms, including Genetic Algorithm and Differential Evolution, to fine-tune weight adjustments, resulting in a substantial enhancement of overall classification accuracy.

Database and Business Intelligence for Mumsnet Dataset | Oct. 2023 – Dec. 2023

Tools Used: DBMS, SQL, Visual Studio, Microsoft SQL Server 2018 Analysis Services

- Executed database management and business intelligence projects for the Mumsnet dataset, focusing on data normalization, stored procedure development, and business intelligence platform creation.
- Normalized Order Item, Product, and Customer tables to 3NF, enhancing data integrity. Crafted robust stored procedures for order creation and item addition with error handling.
- Spearheaded development of a robust business intelligence platform utilizing Microsoft SQL Server 2018 Analysis Services.
- Designed intricate cube measures and dimensions to comprehensively capture essential business metrics, contributing significantly to informed and strategic decision-making processes.
- Delivered detailed breakdowns of sales and orders, meticulously categorizing data by product, and customer.

E. coli Protein Localization Prediction Using Machine Learning

Tools Used: Python, Pandas, Scikit-learn, Matplotlib, Seaborn.

- Developed Python models leveraging Scikit-learn to predict E. coli protein localization, employing Gaussian Naïve Bayes and Logistic Regression for classification.
- Compared model performances using metrics like accuracy, precision, recall, and F1-score, achieving superior prediction accuracy with the Logistic Regression model.
- Conducted data pre-processing and analysis on the UCI Machine Learning Repository dataset, showcasing proficiency in Pandas for data manipulation and Matplotlib/Seaborn for visualizations.
- Demonstrated skills in machine learning, statistical analysis, and data visualization, with detailed project information.

Loan Sanction Prediction Using Machine Learning in R

Tools Used: R programming Language, Packages (tidyverse, corrplot, caret, rpart & rpart. Plot)

- Conducted an in-depth EDA on a loan dataset in R, focusing on key variables such as Age, Income, and Credit Score.
- Applied data pre-processing techniques, including handling missing values & encoding categorical variables, to prepare the dataset for ML modelling.
- Developed and compared two machine learning models, Linear Regression and Decision Tree, for predicting loan sanction amounts, and evaluating model performance using Mean Squared Error (MSE).
- Visualized model results using histograms, boxplots, scatter plots, and correlation heatmaps, providing insights into the loan sanction prediction project.