SAURAV BISHT | Data Scientist

United Kingdom | (+44) 7766702629 | saurav.bisht1709@gmail.com |
https://www.linkedin.com/in/saurav-bisht02/ | https://github.com/sauravbisht17

EDUCATION

Master of Science in Data Science   January 2023 – January 2024
University of Greenwich   *London, United Kingdom*

Bachelor of Technology in Information Technology   August 2014 – June 2018
G. L. Bajaj Institute of Technology and Management   *Uttar Pradesh, India*

WORK EXPERIENCE

Associate System Engineer   July 2018 – August 2022
IBM India Pt. Ltd.   *Pune, India*

Nedbank API Marketplace Project:

- Led customer segmentation analysis as an analytical problem-solver of over 1 million transaction records using Python's Pandas and Scikit-Learn's clustering algorithms to enable investigative approach of 10+ targeted marketing campaigns, increasing customer engagement by 25%.
- Developed and implemented risk analysis model, showcasing strong critical attention to detail using logistic regression and random forests from Scikit-Learn to assess creditworthiness of 500,000+ customers, reducing default rates by 15% and enhancing loan approval process.
- Conducted customer churn analysis across 200,000 accounts using survival analysis, gradient boosting models from Scikit-Learn and XGBoost, identifying 5 key factors through an innovative mindset that decreased annual churn rates by 10%.

Vodafone Idea Limited Project:

- Performed customer behaviour analysis, developed customized, customer-focused strategies on 2 million+ subscribers using association rule mining and market basket analysis to develop 20 customized marketing strategies, increasing customer satisfaction scores by 30%.
- Analysed 50,000+ customer complaints using NLP techniques like topic modelling and sentiment analysis from NLTK and spaCy, identifying 3 major issues and prompting effective operational changes that reduced complaint rates by 40%.
- Applied predictive analytics in a data-driven manner using time series forecasting, neural networks from TensorFlow and Keras to assess network traffic for 10 million subscribers, guiding $20M infrastructure expansion that improved capacity by 50% and reduced congestion by 35%.

PROJECTS

Stock Sentiment Analysis   January 2024 - April 2024
Data Scientist   *London, United Kingdom*

- Developed an end-to-end machine learning pipeline showcasing technical expertise for sentiment analysis on stock-related text data, including data cleaning, preprocessing, exploratory analysis, and model building with custom deep neural network architectures using TensorFlow and Keras, incorporating inventive methods with Embedding layers, LSTM (Long Short-Term Memory) layers, Dense layers, and Dropout layers for sentiment classification tasks.
- Utilized Python data analysis libraries (pandas, numpy, matplotlib, seaborn) and natural language processing libraries (wordcloud, nltk, gensim) for data manipulation, visualization, and text preprocessing tasks.
- Leveraged the state-of-the-art BERT model from the Hugging Face Transformers library, fine-tuning it on the stock sentiment dataset and demonstrating inference capabilities on new, unseen text data for sentiment prediction.

Data Analysis for Retail Company   January 2023 - January 2024
Data Analyst   *London, United Kingdom*

- Conducted a comprehensive data analysis, creating 8 impactful visualizations through meticulous data storytelling for a retail chain across 45 outlets using Python. Uncovered trends that led to a 15% increase in customer retention and 10% sales growth, significantly influencing marketing and operational strategies.

Road Accident Severity Prediction Project for UK   January 2023 - January 2024
Machine Learning Engineer   London, *United Kingdom*

- Leveraged machine learning and neural network techniques with a problem-solving mindset to analyse over 150,000 UK road accident records from 2019, delivering high accuracy rate of 95% in predicting accident severity levels. Authored a comprehensive report comparing the efficacy of traditional algorithms versus neural networks in predictive accuracy.

CERTIFICATION

- Oracle Certified Infrastructure: Generative AI Professional
- Microsoft Certified: Azure Data Engineer Associate
- Databricks Certified: Data Engineer Associate
- AWS Certified Cloud Practitioner
- IBM Cloud Private Infrastructure and Architecture
- IBM Certified Solution Implementer - WebSphere DataPower SOA Appliances, Firmware
- CompTIA Linux+: Introduction to Linux & the Command Line
- Forage - Data Visualisation: Empowering Business with Effective Insights
- Analytics Vidhya - Machine Learning Certification Course for Beginners

SKILLS

- **Programming Languages:** Python, R, SQL, CQL.
- **Machine Learning:** Regression, Classification, Clustering, Neural Networks, Natural Language Processing.
- **Data Analysis Tools:** Pandas, NumPy, Scikit-Learn, TensorFlow, Keras.
- **Data Visualization:** Matplotlib, Seaborn, Plotly, Tableau, Power BI.
- **Big Data Technologies:** Hadoop, Spark, Apache Kafka, Apache Pig, HBase, Hive.
- **Database Management**: MySQL, PostgreSQL, MongoDB, AllegroGraph, Neo4J, Cassandra, Redis.
- **Cloud Platforms:** AWS, Azure, Google Cloud.