Vikash Goyal

Senior Data Scientist, 7+ Years' Experience

Email: vikashgoyal2701@gmail.com   LinkedIn: www.linkedin.com/in/vikash-goyal   Phone: +91-9760173419

PROFESSIONAL SUMMARY

Data Science lead at Citi with **7+** years of experience in business decision making by using **Machine learning, Deep learning algorithms** & **Artificial Intelligence**. Currently leading projects in **Customer Experience** vertical. Responsible for requirement gathering, development & delivery of end-to-end solutions. Industry exposure in **Banking, Retail, Healthcare** etc.

EXPERIENCE

CITI | LEAD DATA SCIENTIST
February 2020 – Current | Bangalore, India

→ Developed an **Early warning detection system** to predict hidden attrition from institutional clients
- Hidden attrition reduces **40%** of revenue growth. Designed a multi-dimensional framework to prioritise customers with high attrition risk and low regulatory risk
- Worked in a cross-functional role with different product & risk teams to leverage datasets like **Transactions, Balances, Pricing, VOC surveys, Call reports & Deals**.
- Trained (**XGBoost** model giving an accuracy of **82%**(covering **292** clients from **356** clients provided by **C@R**) and captured **$140 Mn** at risk

→ Developed a **Shell Company Account identification model** using **Network Analysis**
- Worked with business, vendors & ingestion team to expose third party vendor data
- Used **Community Detection Algorithm** to isolate suspicious accounts. Trained **Isolation Forest** model using features from transaction feed & company profile to detect anomalies.
- Model is working at **76% accuracy** in detecting potential shell accounts. Led to **3 FTE reduction** from AML research team

→ Automated the transaction monitoring process which generates insights for **KYC renewal** of accounts. As a part of that, developed a stacked classifier to tag an entity name as Individual/Corporation using **CNN-1D & GBM**. Accuracy achieved is **96%** and rolled out this algorithm to enhance payments data for **CITI wide consumption**

EXL SERVICES | DATA SCIENTIST
April 2017 – January 2020 | Gurugram, India

→ Developed an **NLP fraud detection model** to identify fraud in text based legal invoices. Features from text data was extracted using techniques like **tf-idf, topic modelling, k-means clustering, Word Embedding etc. and XGBoost** was used to build the final classifier. Captured **82% of frauds in the top 2 deciles** which accounted to **$3+ Mn** potential adjustments after analysing **$40+ Mn** legal spends

→ Developed a binary classifier to predict cricket WC'19 matches for the EXL Hackathon. Built a **web scraper** to extract information around matches, player statistics, ranking etc. from Internet. Various features like **player's historical stats, recent form, pitch impact** etc. were created to train a **Random Forest** model. Achieved an **accuracy of 78%** and secured **4th** rank across EXL

AXTRIA INDIA PVT.LTD. | BUSINESS ANALYST
June 2016 – March 2017 | Gurgaon, India

→ Automated the process used for weekly sales tracking of an Oncology brand. Worked on multiple Ad-hoc analysis to generate insights for a major healthcare firm

EDUCATION

INDIAN INSTITUTE OF TECHNOLOGY (IIT), ROORKEE
B.Tech in Production & Industrial Engineering
July 2012 - June 2016
CGPA : 7.7

SKILLS

PROGRAMMING
- Python
- Pyspark
- Pytorch
- SQL

CORE COMPETENCIES
- Logistic Regression • Random Forest • KNN • GBM • XGBoost • K-Means • Machine Learning• Deep Learning• NLP • Large Language Models LLM • GCP • Topic Modelling • Word Embeddings • Network Analysis • Neural Networks • Isolation Forest • DBSCAN • Hypothesis testing • A/B Testing • Attrition Prediction • AML & Compliance • Fraud Detection

ACCOMPLISHMENTS

- Trained 100+ campus hires in Python, Pyspark & ML
- Recognitions: Platinum Award & Delivery Excellence Award(2022), Silver Award (2021), Copper Award (2020) by AIM Bangalore, Citi
- Cleared CFA level 1 in Jan'18 and CFA level 2 in Aug'18