ABHINAV SINHA
abhinavsinha845@gmail.com • +91 7351928565 • LinkedIn • GitHub • Noida

SUMMARY

Dynamic and results-oriented Data Scientist with more than 4 years of expertise in crafting advanced analytics solutions to tackle business risk challenges. Proficient in converting complex business issues into analytical frameworks, I specialize in building and deploying comprehensive machine learning pipelines using Python and Spark. My skill in statistical analysis enables me to distill vast datasets into actionable insights for strategic decision-making.

SKILLS & INTERESTS

Programming Languages: Python (Pandas, Numpy, ScikitLearn, Seaborn, Matplotlib), SQL
Data Science & Analytics: Apache Spark (PySpark), Machine Learning (Regression, Classification, Clustering, Dimensionality Reduction), Model Development, Model Evaluation, Model Deployment, NLP, Large Language Models, BERT, GPT,LangChain, Tensorflow, Keras, PyTorch, Time series and forecasting, MLOps
Others: GitHub, Databricks, AWS, MS Powerpoint, MS Excel

PROFESSIONAL EXPERIENCE

Dun & Bradstreet India, Chennai, India   March 2023 – Present
Senior Analyst, Data Science
- Instrumental in establishing core scoring rules for ESG ratings, **contributing to $3.2 million in revenue** generation.
- Developed an **NLP-based model** to extract and classify climate risk disclosures from ESG reports, news, and regulatory filings, enabling accurate assessment of regulatory, physical, and market risks for the customers.
- Developed and deployed **LLM**-powered applications using **GPT-4** and **OpenAI API** for extracting insights from company reports, enhancing ESG risk assessment.
- Implemented transformer-based NLP models (**BERT**, GPT-4) leveraging **Hugging Face** libraries to classify ESG disclosures and news articles.

Dun & Bradstreet Technologies and Data Services, Chennai, India   March 2021 – February 2023
Data Scientist II
- Built business default prediction model using XGBoost, leveraging financials, delinquency variables, trade data, and other sources to classify default risk with high accuracy.
- Spearheaded the team charged with **on-boarding and performing data analytics of 1000+ internet data sources** comprising of broad diversity of datasets, increasing the data richness of product, and making it a key component in ESG.
- Developed suite of ML models to predict authenticity of customer's suppliers operating in various markets in Asia and South America, helping the organization secure multi-year contract with Fortune 500 customer. (**Revenue generation: $5 million/year**)

Data Scientist I
- Developed a machine learning model for detecting promo code abuse in a payment app, automating and improving the efficiency of the previous manual process.
- Built time series forecasting model for financial risk prediction, improving default risk classification accuracy for a major European bank customer.

KPIT Technologies, Pune, India.   June 2018 – May 2020
Software Engineer
- Utilized Python for CAD modeling and automation in the design and development of wire harness systems for the electrical architecture of a next-generation concept super truck, enhancing design efficiency and accuracy.

INTERNSHIP

Dun & Bradstreet Technologies and Data Services, Chennai, India   Nov 2020 – Feb 2021
Data Science Intern
- Developed Python and PySpark tools that significantly expedited model development, now integral to major ML projects in the organization.
- Migrated data transformation and modeling pipelines in Pandas to PySpark, thereby **improving end-to-end process runtime 10x**.

EDUCATION

University of Petroleum & Energy Studies | Dehradun, India | B.Tech Electrical Engineering   May 2018