

Udit Bhola

Senior Manager | Lead Data Scientist

Get in Touch @ | 9888701925 | bpgp24udit@iima.ac.in
in [linkedin.com/in/uditbhola](https://www.linkedin.com/in/uditbhola)

Industry Experience

Lead Data Scientist - M&G plc

21st January 2019 - present

Key Roles: Leading Ideation and Development of AI/ML Solutions across Private and Public Assets for Portfolios, Investment Research, Sustainable Investing. Demonstrating proficiency in navigating diverse cultural and business environments with significant international exposure in London and Vietnam, fostering global collaborations and cross-cultural understanding.

Key Deliverables:

- **Developed AI-enabled solutions for early adoption to enhance efficiency and optimize existing processes.**
Leveraged Azure OpenAI, LangChain, LlamaIndex, Vector DB, open-source LLMs, and embedding models to create applications such as a RAG-based information retrieval tool, Excel Smart Insights, Investment 360°, an AI-enabled presentation generator, and Text2SQL. Employed cutting-edge chunking strategies, fine-tuned embedding models, and query decomposition techniques with a reranker to improve information retrieval and accuracy.
- **Scope 1/2 and Scope 3 Prediction Tool for Private Assets, RMBS and Social Housing Emission Estimation.**
Streamlit based Web App Deployed on Azure Devops with ML models serving on Databricks API's. Solution Solved the problem of no emission disclosures of Private Assets worth £7.3billion aiding into better regulatory reporting and Responsible investing aligned with Net Zero Targets of M@G plc.
* This tool has also been cited in M&G plc's Sustainability Report.
- **AI Enabled Smart Extraction for Unstructured Documents.**
Intelligent Extraction using CNN to extract and parse only financial tables from any and all pdf's and extracting structured contents from any pdf's into excel spreadsheets. Also, generating an editable pdf from scanned pdf's to allow analysts to use search and find operations.
The internal tool replaces the Adobe licensed tool to convert scanned/image based pdf's to text based pdf's.
- **Automated PDF Parser for Fund Reports using a Combination of Azure Recognition, Tabula, Pdf Plumber and Camelot to Parse 200+ Fund Reports Monthly to extract critical Financial Metrics and Tables(Databricks Scheduled Jobs along Azure Logic Apps were used to deliver the self serviced Solution to End User)**
Extracting the PDF's Content into Predefined Excel Templates saved 800+ hours of Resource timings along with reducing the TAT for validation and delivering benefits worth £18,188 cost saving.
- **BERT Based ESG Scoring to create Carbon Intent Scores parsing Annual Reports and Earnings Calls Quarterly Reports(~50K) to generate Climate Transition Scores for Organizations using FactSet data.**
Solution Aided into Investment Research quantifying the Climate Transition Actions into Climate Score Undertaken by Organizations. Contributing as an integral Score for ESG Scoring.
- **Web scraping Websites (re100,tcf,CSR, responsibility reports) to fetch the sustainable Net Zero goals for Index organizations.**
NetZero Targets From Alternate Data Sources Helped to Keep the ESG investing Portfolios updated reducing the TAT compared with other annual data sources(e.g. CDP). Also, these targets were used to fill internal scorecards.
- **Developed and Published Analytical ESG, Climate and ABS Power BI Dashboards using NLP and ML under the hood, with End to End ETL Pipelines from MSCI, Bloomberg, CDP and ISS Source Data Feeds.**
These Dashboards are widely used to investigate, research and invest into ESG pool companies. The Dashboards contain granular level ESG data Segregated into Climate, Social and Governance Sections data.
- **Extract Stock Sentiments using GDELT and Lexis Nexis data and apply Financial and ESG Sentiment analysis.**
Using Real Time News Data to understand the Investment Themes trending in the Digital World to aid into tracking of Interest shifts across geographies.
- **Using regression discontinuity to Analyze debts worth collecting for Leveraged Loans Team.**
Loan Books of Acquired portfolios are analyzed for NAV metrics for better Decision makings.
- **Anomaly Detection Pipelines for ESG data feeds to identify False Reporting and Green washing.**

Objective:

An Analytics professional with skills in Data Science, Business Analytics, Product Management & Big Data with 8+ years of experience in Leveraging Analytics to create Solutions and enable businesses to Achieve their growth goals and ambitions.

Text section

Data Scientist with ability to use industry knowledge, contextual understanding, skepticism of existing assumptions – to uncover solutions to business challenges.

Skills

Generative AI



Expert in leveraging Generative AI techniques to drive innovative data solutions and predictive modeling, enhancing business decision-making and operational efficiency.

Python



Well Versed with all Data Science libraries like pandas, Numpy, scikit learn and matplotlib along with practical implementation knowledge of deploying ML models using Azure Devops Stack.

Machine Learning



Supervised and Unsupervised Machine Learning techniques.

Rule Based and ML based logics are devised to identify change in business model or Green washing.

- **Twitter Based Data Insights for FTSE100 CEO's Public Accounts.**
CEO's post are used as a proxy to validate the Social and Governance Scores of the organizations collating with the Indeed Company ratings data.

Senior Data Scientist

15th October 2018 - 18th January 2019

[A.P Moller - Maersk -Mumbai](#)

Key Roles:

Project Coordinator of Data Science team for Mumbai Region.

- Providing Analytical Solutions to Sales, Marketing, Operations and Finance Departments for improving KPI' s in terms of Volumes and Sales along with making the customer on boarding journey smoother

Data Scientist

August 2017 - - September 2018

[HDFC Life- Mumbai](#)

Key Roles:

- Prediction Model for Propensity to give an appointment for cross selling based on customer interaction data and pushing the results to Apache Kafka.(Led to Reduction of calling cost upto 1 cr per quarter)
- Policy Surrender Fraud Detection (Reduction in Frauds by 7 % after implementation of the model).
- Facial recognition system for Existing Customers using Neural Networks and KNN.
- Churn Modelling for Policy Surrender Using Artificial neural network.
- Prediction Model to Predict Premium and Total cost of Ownership of a Policy Using XGboost and deploying the model with a Rest API.
- Developed a Model to Solve the Problem of Deduplication and Deployed the model using REST API.
- Developed an Address Parser for Internal team to get Geo coordinated addresses for Indian Addresses using R.
- Customer Segmentation for Recommending Products using Clustering.
- Designing Income Surrogates using Decision Tree from Alternate Data Sources.
- Product Owner : Ideated, created and delivered suite of analytical solutions for products to multiple stakeholders in organization.
- Working Closely with Senior Management of Different Departments to take requirement for analytical solutions and leading the product development life cycle related to Data Science.

Decision Analyst

August 2015 - - September 2016

[BSS Engineering Services -PAN India](#)

Key Roles:

- Worked on off roll basis and was absorbed as a permanent employee after 1 year.
- Predicting energy output of a Combined Power Plant.
- Predictive Production planning, execution and variance analysis with impact study on Power plant performance.
- Planning for preventive & predictive maintenance of machinery using Machine Learning Ensuring maximum machine uptime.
- Power BI Dashboard Reporting for fuel consumption and Resource allocation.

EDUCATION:

Blended Post Graduate Programme(MBA)

2024 - 2026

[Indian Institute of Management, Ahmedabad](#)

PGP in Data Science, Business Analytics and Big Data.

2016 - 2017

[Aegis School of Business and Data Science \(Mumbai\) jointly certified by IBM.](#)

IBM Certified Data Science Professional

Post Graduate Diploma in Operations Research Management(DEC-UGC-AICTE)

2015 - 2016

[Jaipur National University](#)

Deep Learning



Application of ANN, LSTM and CNN Architectures for Real Business Solutions

Agile Methodologies



Experience of Working in Agile methodology following Scrum Software development method with tools Jira and Confluence.

Statistics



Descriptive and inferential statistic methods along with statistics of Machine Learning.

Apache Spark



Working with Big Data using Amazon Emr and Delta tables on Databricks

DataBricks



Using Databricks to build ML models and Poc's

Power BI



Interactive Dashboards, Data Visualisations and Complex Calculations using Python Streamlit and Power BI

SQL



Certification from Solo Learn

B. Technology(UGC)

2010 - 2014

MONAD UNIVERSITY

Electrical Engineering

Data Science Community Activities:

Industry Expert - Data Science & Analytics

August 2023 - August 2024

[Upgrad Campus](#)

Played a key role as SME in the development and delivery of Flagship Data Science programs for Upgrad Campus.

Subject Matter Expert

January 2020 - January 2021

[WhitePanda](#)

Creating Analytical Training Content for Business houses.

Subject Matter Expert(Application of Data Sciences and Analytics in Financial Services)

April 2019 - January 2020

[Mumbai University](#)

On Panel for SME'S working in Financial Services.

Mentoring Students on career transitions and workshops using real world use cases in financial sector for leveraging Data Sciences

Participant in Analytics Vidhya Hackathons:

Top 10 percentile Score.

Achievements:

- Consecutively Top Rated Employee in Data Science Team Since last 5 years.
- Got Multiple Recognitions for AI projects built for the Prudential Investment Business.
- Receiver of Special Recognition, Achievers Award, Multiple Star of the Months and Icon of the year awards.
- Internal Hackathons Runner up Twice.
- Engaging with High Seniority Stakeholders and Fund Managers.

