

# ADITYA MISHRA

## Data Analyst | Machine Learning Engineer | Data Scientist

[LinkedIn](#) | +91 7007788926 | [portfolio-website](#) | adityamishra0996@gmail.com | [GitHub](#)

### Skills

- Python, R, Machine Learning, SQL, Git, Amazon Redshift, Power BI, Zoho Analytics, Google Analytics, ETL Scripting, Excel
- NumPy, Pandas, Matplotlib, Seaborn, Scikit-learn/Sklearn, NLTK, Genism, Streamlit, Bertopic, UMAP, HDBSCAN, Pickle
- Data Gathering, Data Mining, Data Cleaning, Data Manipulation, Data Visualization, Data Modeling, Data Analytics
- Statistics, Statistical Analysis, Statistical Modeling, Predictive Analytics, Predictive Modeling, Natural Language Processing
- Business Analysis, Business Insights, Business Acumen, Knowledge Acquisition, Presentation Skills, Storytelling, Critical Thinking
- Generative AI, LLM(s), Basics of BigQuery, Prompt Engineering

### Professional Experience (2.7 Years Relevant)

#### Business Analyst

[Excalibur Technologies Pvt. Ltd.](#)

Gurugram, India

01/2021 – 02/2022

- Gathered and analyzed business requirements through collaboration with stakeholders, ensuring alignment with project goals.
- Prepared comprehensive documentation including BRD, FRD, and Use Case Specifications.
- Performed data analysis to support informed business decisions and maintained thorough documentation of project developments and changes.

#### Data Analyst

[Edureka](#)

Bengaluru, India

02/2022 – 09/2023

- Produced monthly reports using advanced Excel functions, driving data-driven decisions.
- Increased sales agent productivity by 10% by creating and maintaining sales metrics files.
- Reduced incentive calculation time by 50% by automating the process.
- Optimized **Power BI** dashboards, achieving up to 95% size reduction, leading to improved efficiency, resulting in faster load times and reduced resource consumption.
- Developed a **Course Recommendation System** that increased learner engagement by 20%, delivering the top 5 course suggestions. Employed cutting-edge libraries and techniques, including Scikit-learn, SQL, NLTK, and Streamlit, alongside TFIDF Vectorizer, Bag of Words, and PorterStemmer for enhanced recommendation accuracy.
- Implemented a QA Posts Sentiment Analysis System that enhanced learner satisfaction by reducing support tickets by 20%. Leveraged advanced libraries and techniques like Sklearn and NLP for thorough Exploratory Data Analysis, Feature Engineering, and precise sentiment analysis. Utilized tools such as Word Tokenizer, Word Lemmatizer, TextBlob's Polarity, and topic modeling techniques like Latent Dirichlet Allocation (LDA) and Non-Negative Matrix Factorization (NMF) for comprehensive insights.
- Built a topic model (BERTopic Model) to pinpoint key learner behavior patterns, facilitating targeted enhancements in critical areas. This strategic approach led to increased learner satisfaction and revenue growth. Employed a sophisticated toolkit, including libraries like Sklearn, NLP, NLTK, Genism, and Sentence-Transformers, and techniques such as CTF-IDF, UMAP, HDBSCAN, LDA, and NMF, to extract actionable insights and guide organizational improvements.
- Resolved 500+ Jira Cards, ensuring streamlined operations.
- Developed over 100+ Power BI & Zoho Analytics Dashboards from scratch, empowering data-driven insights.
- Maintained Key Performance Indicators, ensuring accurate tracking of performance metrics.
- Created a Power BI Golden Dataset for B2C Business Vertical, which includes more than 20 dashboards, over 160 redshift tables/views, and DAX () tables, along with Google Sheets.

### Education

#### Bachelor of Technology

[AKTU](#)

Ghaziabad, India

08/2015 - 08/2019

- Major in Computer Science Engineering

### Projects

- Machine Learning Dataset Explorer:** Created a simple ML dataset explorer app for Data and visual level EDA (Python, Streamlit). Link to [GitHub](#) Repository (12/2022)
- Hate2Wait:** Designed and developed an award-winning website for the State Level Innovation Competition in Ghaziabad (HTML, CSS, PHP) (12/2019)
- Healthcare NIDDK:** Built a model to accurately predict whether the patients in the National Institute of Diabetes and Digestive and Kidney Diseases dataset have diabetes or not (Python, k-Nearest Neighbors, AUC/ROC, Tableau) (06/2022)
- Snake Game:** A Reinforcement Learning game, which learns automatically how to play the Game by Maximizing the notion of Reward. (Python, PyTorch, Deep Learning). Link to [GitHub](#) (01/2023)

### Certifications

- Certification of Excellence:** Won State Level Innovation Competition, held in Ghaziabad. (12/2019)
- Certification of Achievement - Data Scientist Masters Program:** Simplilearn, Link to [Certificate & Creds](#) (06/2022)
- Machine Learning Advanced Certification Training:** Simplilearn, Link to [Certificate & Creds](#) (11/2021)
- Data Science Capstone - Healthcare:** Simplilearn, Link to [Certificate & Creds](#) (06/2022)
- Data Science with Python:** Simplilearn, Link to [Certificate & Creds](#) (09/2021)
- Data Science with R:** Simplilearn, Link to [Certificate & Creds](#) (07/2021)
- SQL Training:** Simplilearn, Link to [Certificate & Creds](#) (03/2021)
- Git Training:** Simplilearn, Link to [Certificate & Creds](#) (07/2021)

### Languages

- English:** Proficient
- Hindi:** Native/Bilingual