Vanshita Arya

+91-969-422-2811 | vanshita.arya@outlook.com | [LinkedIn](#) | [GitHub](#) | [Kaggle](#)

EDUCATION

Amity University   Noida, U.P.
Masters in Data Science   *Aug. 2023 – Present*

Banasthali Vidyapith   Tonk, Rajasthan
Bachelors in Maths, Statistics and Physics   *Aug. 2017 – May 2020*

EXPERIENCE

Data Science Intern   March 2024
British Airways

- Led a simulation highlighting data science's importance at British Airways, resulting in a **20%** simulated revenue increase.
- Used advanced web scraping to analyze 10,000 customer reviews, boosting satisfaction ratings by **15%**.
- Developed an **85%** accurate predictive model, enhancing marketing effectiveness by **25%** through insights into buying behaviors.

Social Media Analyst   December 2023 – Present
dsaa2z

- Leveraged data-driven insights to increase engagement and follower growth on instagram, achieving a **+10.9%** increase in accounts reached and a **+24.1%** increase in accounts engaged over the last 30 days.
- Successfully grew the channel to **201K** followers, reflecting a **+6.8%** increase from the previous month.
- Provided actionable recommendations for content planning, focusing on aligning with audience needs and preferences to sustain consistent growth and engagement.

PROJECTS

National Stock Exchange Daily Performance Data| *Kaggle API, datetime module, NSE Bhavcopy*   July 2024

- Published a comprehensive dataset on Kaggle featuring performance data of top NSE scripts from January 1, 2024, updated weekly.
- Achieved a usability score of 8.82, collaborated with data science community on LinkedIn, Kaggle, GitHub.
- Integrated data into a single Parquet format for efficient storage and querying.

Amazon Fashion Sentiment | *Google Colab, NLP Modules, Hugging Face, NLTK*   June 2024

- Conducted random sampling of 10,000 Amazon fashion customer reviews spanning 1993 to 2023.
- Performed exploratory data analysis revealing over 57 % of reviews are positive, with few negative reviews. .
- Applied NLP models like VADER and fine-tuned RoBERTa, correlating sentiment scores with star ratings, with RoBERTa capturing finer sentiment nuances .

TECHNICAL SKILLS

Languages: Python, Excel, R, MySQL
Frameworks: NumPy, Pandas, Scikit-learn, Matplotlib, Seaborn, Tensorflow
Developer Tools: Git, Azure, Google Colab, VS Code, Jupyter Notebook

ACHIEVEMENT

Dataset Publication| *Kaggle* (2024): Published a comprehensive dataset titled "National Stock Exchange Daily Performance Data" on Kaggle, recognized with a DOI: 10.34740/KAGGLE/DSV/9026608.

Leadership Roles| Currently serving as Placement Coordinator (2024-Present), demonstrating enhanced leadership and communication skills. Served as Class Representative (2023-2024), coordinating with peers, resolving conflicts, and collaborating with other representatives. Also acted as Prefect of the hostel, strengthening resilience.