Yash Vyas

Phone: +19309041801 | Mail: yashvyas@iu.edu | LinkedIn: linkedin.com/in/yashhvyass
GitHub: github.com/yashhvyass | Portfolio: yashhvyass.com

EDUCATION

Master of Science in Data Science   Aug 2024 - Expected May 2026
Indiana University Bloomington, Luddy School of Informatics, Computing, and Engineering
Relevant Coursework: Statistics, Probability, Applied Machine Learning, Advanced Database Concepts, Deep Learning

Bachelor of Science in Computer Science   Aug 2017 - May 2021
Vellore Institute of Technology   GPA : 3.7/4
Relevant Coursework: Data Structures and Algorithms, Linear Algebra, Computer Networks, Data Analysis

EXPERIENCE

Data Scientist   July 2021 - July 2024
Signify (Philips)   *Bangalore, India*

- **Recognized as Top Performer** for optimizing ROI by **10%** across 6 global markets applying Market Mix Modeling and designed a KPI dashboard with Streamlit, reducing execution time by **2 weeks**
- Reduced Global Consumer Channel attrition by **6%** through weighted gradient boosting, NLP for text analytics, OCR for extracting information from resumes (text, images), and DBSCAN for employee segmentation
- Implemented topic modeling and sentiment analysis on scraped customer reviews utilizing Beautiful Soup, Snowflake, and BERT, contributing to **$2 million** increase in revenue for targeted products
- Orchestrated a Power BI simulation tool for price elasticity analysis, influencing strategic pricing adjustments
- Co-developed a hybrid recommendation system to provide product suggestions leveraging ALS algorithm, Snowflake, and AWS, boosting multi-device purchases by **25%**
- Collaborated with data engineers to automate data validation for MMM, **saving 1 week** in data collection

Data Science Engineer Intern   Feb 2021 - June 2021
Signify (Philips)   *Bangalore, India*

- Integrated R functionalities into Python by adopting Rpy2, improving cross-functional project workflows
- Built time series forecasting pipelines for financial planning across **3 markets**, applying ARIMA, XGBoost, LSTM, and 1D CNN, enhancing accuracy in market predictions

Data Risk Analytics Intern   May 2020 - June 2020
KPMG   *Remote*

- Devised credit risk models using SVM to predict loan defaults for clients with limited credit history leveraging Gini metrics

PROJECTS

FarmEasy | Large Language Models (LLM): Link

- **Finalist** in Agrinovus Producer-Led Innovation Challenge at IUB among **50+** candidates involving companies and students
- Developed a multimodal RAG platform with Mistral, Google API , Whisper, Streamlit, and JavaScript, enabling farmers to expand market reach across states with real-time, accessible guidance on regulations, pricing, and logistics

Textly | Natural Language Processing, PyTorch: Link

- Developed Textly, an advanced alternative to Quillbot, employing Flask and a fine-tuned BERT model for style-specific, multilingual support, achieving a BLEU score of **45**, and deployed with Docker and Vercel.

No Brainer | Computer Vision, Medical Image Classification, Tensorflow: Link

- Created a real-time Flask app to classify brain tumors using transfer learning on ResNet50, VGG, attaining **97% f1 score**

LEADERSHIP AND HONORS

- Facebook Developer Circle: **Campus Representative**, facilitated online events for **100+** students on python
- Leaf Classification Challenge: Achieved **Top 6%** in MachineHack's Computer Vision Hackathon

SKILLS

Programming and Development:	Python, R, Matlab, SQL, PySpark, Pytorch, Tensorflow, Docker, AB Testing
Cloud and Data Platforms:	AWS, S3, Google Cloud, SAP, Snowflake, Tableau, Power BI, Streamlit