

YASH VYAS

Mail: yashhvyass@gmail.com ◇ LinkedIn: [linkedin.com/in/yashhvyass](https://www.linkedin.com/in/yashhvyass) ◇ Github: github.com/yashhvyass

EDUCATION

Masters of Data Science, Indiana University Bloomington Aug 2024 - May 2026

Bachelor of Computer Science, Vellore Institute of Technology (GPA:8.4/10) Aug 2017 - May 2021

Coursework: Statistics, Probability, Data Structures, OOPs, Machine Learning, Deep Learning, Big Data Systems

SKILLS

Languages and Platforms Python, R, SQL, AWS, SAP DI, GCP, Tableau, Power BI, Snowflake, Streamlit
Other Docker, Flask, FastAPI, AB Testing, Pandas, sklearn, Pytorch, LLMs, GPT

EXPERIENCE

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

- Implemented Market Mix Modelling, optimizing ROI by an average of 10% across 6 global markets using UCM algorithm and designed an intuitive KPI monitoring dashboard with Streamlit, to reduce execution time by 50%
- Reduced attrition by 6% for the Consumer Channel by using weighted SMOTE LightGBM, achieving an impressive 90% F1 score on employee and resume data extracted using NLP and OCR techniques through a well-structured ETL pipeline
- Developed a sophisticated regression-based simulation tool, incorporating appropriate transformations for price and sales optimization; achieved a notable 15% increase in market share and \$500k in annual revenue growth
- Leveraged Snowflake, cutting-edge NLP, and the BERT transformer for text analytics and sentiment analysis on scraped customer reviews, leading to a remarkable \$5 million revenue growth for selected products
- Automated the data validation process for MMM by creating an effective stored procedure in Snowflake streamlining communication between stakeholders, reducing time by almost 30%

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

- Integrated R-based models into Python using the Rpy2 library, which enabled fellow team members to leverage R's capabilities within a Python environment
- Formulated robust forecast models for the Financial Planning and Analysis team by utilizing traditional ARIMA-based time series, machine learning algorithms, and LSTM networks for improved accuracy

PROJECTS

Textly (NLP, Transformers, Pytorch, Docker, FastAPI, AWS, Git) ([Try it here](#)):

- Built an end-to-end text paraphrasing application using Hugging Face's BERT, achieving a BLEU score of 45, offering various paraphrasing styles, reducing text processing time to under 5 seconds per request

EduStar (LLM, Llama, GPT, Langchain, Flask, Git) ([Try it here](#)):

- Developed a narrative-based learning system using a dual-interface chatbot for teachers and students, incorporating LangChain-supported LLM (llama) increasing student engagement and comprehension

HONORS

- **Top performer** at Signify (Philips Lighting): MMM
- **Top 10%** in Leaf Classification Challenge (CNN): Image Classification Hackathon conducted on MachineHack
- **Top 10%** in Analytics AI Challenge: Mutli-class Classification hackathon conducted on Dockship.io