# HARSH RANA

+49 176 8592 7737 ◇ Dortmund, NRW
harshgrana1997@gmail.com ◇ linkedin.com/harshweb98 ◇ github.com/Ianalysebetter

## OBJECTIVE

Data Science Master's student at TU Dortmund with expertise in collecting, processing, and analyzing large and complex datasets. Proficient in statistics, machine learning, AWS, and natural language processing (NLP). Actively seeking a student job or internship in Data Science or Data Analytics to leverage my technical skills and contribute to impactful projects.

## EDUCATION

**Master of Data Science**, Technical University Dortmund — *Expected 2025*
**Bachelor of Information Technology(1.3 GPA)**, Gujarat Technological University — *2015 - 2019*

## SKILLS

| | |
|---|---|
| **Programming Languages** | **Python, R, SQL**, Java, C# |
| **Data Science** | Data Exploration & Transformation(ETL), Machine Learning Model Training, Hyper-parameter Tuning, Statistical Modeling, Hypothesis Testing, Natural Language Processing (NLP), Text Summarization, Sentiment Analysis, Deep Learning, Image Classification, Neural Networks, TensorFlow, PyTorch |
| **Tools and Technologies** | Jupyter Notebook, RStudio, Power BI, Tableau, AWS, Git, Docker, CI/CD (GitHub Actions), Streamlit |
| **Languages** | **English (B2), German (A2), Hindi (Native)** |

## PROJECTS

**1. Deep Learning Project for Kidney Disease Classification** (View)

- **Project Overview:** Developed a comprehensive deep learning model to classify kidney disease, integrating advanced MLOps tools for efficient workflow management and deployment.
- **Data Management:** Utilized **Data Version Control (DVC)** to handle datasets, ensuring reproducibility and effective tracking of data changes throughout the project lifecycle.
- **Model Development:** Implemented a deep learning architecture tailored for medical data, focusing on accurate classification of kidney disease.
- **Experiment Tracking:** Employed **MLflow** to monitor experiments, facilitating comparison of model performance metrics and hyperparameter tuning.
- **Continuous Integration/Continuous Deployment (CI/CD):** Established automated pipelines for testing and deployment, enhancing the reliability and efficiency of model updates.
- **Deployment:** Deployed the trained model to a production environment, enabling real-time predictions and integration with existing healthcare systems.
- **Outcome:** Delivered a robust and scalable solution for kidney disease classification, demonstrating proficiency in deep learning, data versioning, experiment tracking, and deployment practices.

**2. Deep Learning Project Using MLOPS DVC Pipeline With Deployments on Azure and AWS** (View)

- **Project Overview:** Developed and deployed a deep learning model for image classification, implementing MLOps practices to ensure efficient and reproducible workflows.
- **Data Version Control:** Utilized **DVC (Data Version Control)** to manage datasets and model versions, facilitating collaboration and experiment tracking.

- **Continuous Integration/Continuous Deployment (CI/CD):** Set up automated pipelines using **GitHub Actions** to streamline testing, validation, and deployment processes.
- **Cloud Deployment:** Deployed the model on **Azure and AWS**, ensuring scalability and accessibility.
- **Outcome:** Achieved a robust and scalable deep learning solution with automated workflows, enhancing efficiency and collaboration.

**3. NLP-Based Text Summarization Implementation**(View)

- **Project Overview:** Developed a natural language processing (NLP) model for automatic text summarization, utilizing **transformer models** such as **BERT** and **GPT**.
- **Techniques Implemented:** Applied **extractive** and **abstractive summarization** to generate concise summaries from large text datasets.
- **Tools Used:** Leveraged Python libraries such as **spaCy**, **NLTK**, and **transformers** to preprocess text and fine-tune models.
- **Evaluation:** Evaluated model performance using metrics like **ROUGE** and **BLEU**, ensuring high-quality summary generation.
- **Deployment:** Deployed the summarization model as a web application using **Streamlit**, enabling easy access to the tool for end users.

**4. Machine Learning Project Implementation (Boston House Price)** (View)

- **Project Overview:** Developed and deployed a **machine learning model** for predicting **Boston housing prices**.
- **Techniques Implemented:** Preprocessed data, trained the model, and ensured **scalability** through **Docker containerization**.
- **CI/CD Implementation:** Automated **workflows** using **GitHub Actions**, integrating **testing** and **deployment pipelines**.
- **Deployment:** Deployed the model on a **cloud platform**, enabling **real-time predictions** in a **production environment**.
- **Outcome:** Leveraged modern **DevOps practices** to streamline the transition from development to deployment.

## PROFESSIONAL EXPERIENCE

**Data Project Coordinator** *May 2019 – November 2019*
Uplers Private Limited, Ahmedabad, India

- Managed and coordinated multiple projects, ensuring timely delivery and seamless execution.
- Collaborated with international clients to define project scope and align deliverables, utilizing tools like **Jira** and **Confluence** for efficient task management.
- Conducted budget analysis for high-value projects using **Power BI**, enabling cost tracking and proactive risk mitigation.
- Acted as the primary communication liaison, maintaining strong stakeholder engagement through tools like **Slack** and **Skype**, effectively minimizing project delays.

**Junior Data Analyst** *April 2018 – May 2019*
Uplers Private Limited, Ahmedabad, India

- Cleaned and prepared datasets using **Excel** and **Python**, ensuring data accuracy and consistency for analysis.
- Assisted senior analysts in generating detailed reports and visualizations with **Power BI**, aiding clear communication of data insights to business teams.

- Performed routine data entry and extraction tasks using **SQL**, supporting database maintenance and optimization efforts.
- Collaborated on analytics projects, enhancing technical expertise and contributing to successful data processing tasks.

## PROFESSIONAL CERTIFICATION VIEW

- Introduction to Statistics, Stanford University View
- Machine Learning Specialization, Stanford University View
- Practical Data Science Specialization, DeepLearning.ai View
- Microsoft Power BI Desktop for Business Intelligence View