Harikesh Kushwaha

Location: New Delhi, Delhi
LinkedIn | Portfolio | GitHub | Kaggle Email: harikeshkumar0926@gmail.com | Mobile: +919838422934

Data Scientist

I am a graduate student with strong **statistical** background and expertise in **Python** and **Machine learning**. I'm skilled in data analysis and also have a good knowledge of web developement. Morever have strong problem-solving skills and ability to communicate effectively. These skills are useful in my current research and also in my future career and I'm looking forward to apply them in a challenging environment.

Technical Skills

Languages	:	Python (*Proficient*), SQL, JavaScript, C++, MATLAB, HTML, CSS
Frameworks	:	TensorFlow, Keras, Scikit-learn, Django, Streamlit, Bootstrap, react
Libraries	:	matplotlib, plotly, pandas, NumPy, Seaborn, BeautifulSoup, Selenium, NLTK
Databases	:	MySQL, MongoDB
Dev Tools	:	VS Code, Git, GitHub, Jupyter Notebook, Tableau, Excel
Soft Skills	:	Analytical and Problem-Solving Skills, Good Presentation Skills, Quick Learner

Education

Indian Institute of Technology Delhi — New Delhi, India
Master of Science in Physics, (8.6 GPA) — *July 2021 – May 2023 (Expected)*

Banaras Hindu University — Varanasi, Uttar Pradesh India
Bachelor of Science in Physics, (8.4 GPA) — *July 2018 – May 2021*

Projects

Food Vision *Python, TensorFlow, Colab, deployment, Streamlit* Source Code

- Developed a deep **neural network** using TensorFlow and Keras to classify **101 categories of food**.
- Used a pretrained **EfficientNet** model to extract features from the food images, and then **fine-tuned** the model to improve its accuracy.
- Achieved an accuracy of **80%** on the test set, demonstrating the effectiveness of the approach in addressing complex image recognition problems.

NLP With Disaster Tweets *Python, TensorFlow, NLP, Text Vectorization, LSTM, GRU, CNN* Source Code

- Developed NLP models to classify disaster and non-disaster tweets using **text vectorization**, various **word embeddings**, and deep learning models including **LSTM**, **GRU**, their **bidirectional** variants, and **1D CNNs**
- Utilized the **Universal Sentence Encoder** to create embeddings on both the character and word levels, and implemented a **multivariate** model using the **functional API** of **TensorFlow**.

Credit Risk Assesment *Python, pandas, scikit-learn, kaggle*

- Developed a credit textbfrisk assessment model by analyzing various customer features, performing data cleaning, feature engineering, and exploratory data analysis.
- Established a basic model for initial experimentation, and trained advanced models such as **LR**, **SVM**, **XGBoost**, **Catboost**. Top performing model, achieved a test AUC-ROC score of **0.97** and precision of **0.96**.

Certifications

- Machine Learning Specialization (DeepLearning.AI) Certificate
- Deep Learning Specialization (DeepLearning.AI) Certificate
- IBM Data Analyst Capstone Project (IBM) Certificate
- TensorFlow Developer Certificate in 2022: Zero to Mastery (Udemy) Certificate