# Harikesh Kushwaha

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

## Data Scientist

As a recent graduate with a strong foundation in **statistics** and data science, I have worked on several personal projects with real datasets using SQL and Python. In my projects, I have showcased my skills in **data cleaning**, **data visualizations**, and **modeling**. With a passion for solving complex problems and a drive to constantly learn and improve, I am excited to take on new challenges in the field of Data Science.

## Technical Skills

| | | |
|---|---|---|
| **Languages** | : | Python, SQL, JavaScript, MATLAB, C++ |
| **Frameworks** | : | Scikit-learn, TensorFlow, Keras, Django, Streamlit |
| **Libraries** | : | matplotlib, pandas, NumPy, Seaborn, BeautifulSoup, Selenium, OpenCV, Statsmodels |
| **Databases** | : | MySQL, MongoDB |
| **Dev Tools** | : | VS Code, Tableau, Git, GitHub, Jupyter Notebook, Anaconda, AWS, S3 |
| **Soft Skills** | : | Analytical and Problem-Solving Skills, Good Presentation Skills, Communication skills |

## Projects

**House Prices Prediction** *Python, pandas, scikit-learn, kaggle, Matplotlib, Seaborn* Source Code

- Analyzed over **80** features to predict house prices using machine learning.
- Performed **Exploratory Data Analysis** and **feature engineering** to get insight from data.
- Trained **multiple models** using scikit-learn and selected the best one by applying **grid search** and **cross-validation**. Used ensemble of the top performing models to achieve a **top 10%** ranking on Kaggle.

**IBM Data Analytics Capstone Project** *Python, pandas, Matplotlib, Web Scraping, API, Dashboard*

- Gathered and analyzed data from various sources, including **API** and **web scraping**. Conducted **exploratory data analysis** and **wrangling** to prepare the data for further analysis.
- Built a **dynamic dashboard** to extract valuable insights from the collected data, and effectively **communicated** the findings to others through an **engaging presentation**.

**NLP With Disaster Tweets** *Python, TensorFlow, NLP, Text Vectorization, Transfer Learning* Source Code

- Developed Natural Language Processing models to classify disaster and non-disaster tweets using **text vectorization**, **word embeddings**, and deep learning models including **LSTM**, **GRU**, and **1D CNNs**

**Food Vision** *Python, TensorFlow, Colab, Image Processing, Streamlit, Transfer Learning* 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.

## 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 20231*

## Certifications

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