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 machine learning algorithms, I have worked on several personal projects including **sentiment analysis**, **natural language processing**, and **computer vision**. In my recent projects, I have showcased my skills in **data cleaning**, **feature engineering**, and **model selection**. I have also demonstrated my proficiency in tools like **Python**, **TensorFlow**, **Keras**, **scikit-learn**, and **pandas**. 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	: TensorFlow, Keras, Scikit-learn, Django, Streamlit
Libraries	: matplotlib, pandas, NumPy, NLTK, Seaborn, BeautifulSoup, Selenium
Databases	: MySQL, MongoDB
Dev Tools	: VS Code, Git, GitHub, Jupyter Notebook, Anaconda, AWS, Kaggle

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 **data visualization** and **feature engineering** using Matplotlib and Seaborn, respectively.
- Trained **multiple models** using scikit-learn and selected the best one by applying **grid search** and **cross-validation**. Achieved a **top 12%** ranking on the Kaggle leaderboard.

Digit Recognizer   *Python, TensorFlow, Keras, Kaggle*   Source Code

- Developed a very deep **convolutional neural network** using TensorFlow and Keras with **dropout** and **batch normalization** to improve performance.
- Achieved an accuracy of **99.48%** on the test set, securing a place in the **top 15%** on the Kaggle leaderboard.

Food Vision   *Python, TensorFlow, Colab*   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**, 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**.

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, (8.2 GPA)   *July 2018 – May 20231*

CERTIFICATIONS

- Deep Learning Specialization (DeepLearning.AI) Certificate
- Machine Learning Specialization (DeepLearning.AI) Certificate
- TensorFlow Developer Certificate in 2023: Zero to Mastery Certificate