Kumar Ritik

Patna, Bihar | k.ritik8048@gmail.com | +91 7903741636 | LinkedIn | GitHub

Education

Vellore Institute of Technology, Bhopal Sept 2022 – Present
B.Tech in Computer Science and Engineering, Cumulative GPA: 8.23/10

Ishan International Public School, Patna May 2021
Class XII, 74.6%

Ishan International Public School, Patna May 2019
Class X, 92.0%

Experience

Data Science Intern, Sabudh Foundation Jan 2025 - Jun 2025

- Completed modules in Python Programming, Machine Learning, Deep Learning, Natural Language Processing (NLP), Dataiku, and Data Structures Algorithms.
- Collaborated on a team project to implement Deep Learning models to identify brand logos,improving accuracy by 15% over baseline.

Publications

Alzheimer's Disease Detection Using Convolutional Neural Networks (CNNs) In Process
Submitted to Springer.

Projects

Image Captioning and Segmentation Mar 2025 – Jun 2025

- Implemented a deep learning system with ResNet50 and LSTM for captioning and Mask R-CNN for segmentation on MS COCO 2017, achieving BLEU-4 score of 0.35 and IoU of 0.75.
- Optimized NLTK preprocessing and PyTorch data loaders, reducing training loss to 2.32 over 3 epochs.
- Created a Streamlit interface for interactive real-time captioning and segmentation visualization.
- Tech: Python, PyTorch, ResNet50, Mask R-CNN, NLTK, Streamlit, Matplotlib, COCO API.

Alzheimer's Disease Detection Aug 2023 – Nov 2023

- Collaborated in team to build a convolutional neural network using InceptionV3 to classify Alzheimer's stages from 6,400 brain MRI images in a 4-class Kaggle dataset and achieving 94.84% accuracy.
- Optimized data pre-processing pipeline using grayscale conversion and noise reduction, reducing training time by 20%.
- Applied SMOTE for class balancing and data augmentation to enhance model robustness on imbalanced medical data.
- Tech: Python, TensorFlow, Keras, Scikit-Learn, OpenCV, SMOTE, Matplotlib, Seaborn.

Sentiment Preservation Analysis Nov 2024 – Feb 2025

- Developed an NLP pipeline using XLM-RoBERTa to classify sentiment in 5,000 English-Hindi sentence pairs from a 1.66M-pair IITB dataset, achieving 71.40% accuracy.
- Evaluated sentiment preservation, attaining 63.26% accuracy for consistent sentiment between English and Hindi translations on a 2,507-sample test set.
- Implemented robust evaluation with a 0.72 weighted F1-score and confusion matrices, enhancing model performance analysis using Scikit-learn.
- Improved interpretability by 20% through visualizing sentiment distributions and confusion matrices with Seaborn and Matplotlib.
- Tech: Python, XLM-RoBERTa, Scikit-Learn, Pandas, NumPy, Matplotlib, Seaborn.

Technical Skills

- **Programming Languages:** Python, C++, SQL
- **Data Science Tools:** Pandas, NumPy, Matplotlib, Seaborn, Power BI, Tableau, Dataiku
- **Machine Learning:** TensorFlow, Keras, Pytorch, Scikit-Learn Convolutional Neural Networks (CNNs), Long Short-Term Memory (LSTM), Transformers, Natural Language Processing (NLP), Streamlit
- **Computer Vision:** OpenCV, YOLO, Mask R-CNN, Object Detection
- **Databases:** MySQL, SQLite, Oracle

Certifications

- Coursera: Applied Machine Learning in Python
- iamneo: Data Science using Python
- Dataiku: Machine Learning Practitioner, Generative AI Fundamentals

Achievements

- KSP Data-thon 2024 Semi-finalist.
- Completed Deloitte Data Analytics Job Simulation 2025, applying Tableau and Excel for solving real-world tasks.