Tarun Singh

Data Scientist @ Ripik.AI | BTech @ IIT Bhilai

Email: singh.tarun.iit@gmail.com
Mobile: +91-6387610828
https://www.linkedin.com/in/tarun-singh-287a3b199/

EDUCATION

Indian Institute Of Technology (IIT), Bhilai — Chhattisgarh, India
- *Bachelor of Technology - Electrical Engineering* — **Completed** *May 2023*
  CGPA: 8.36/10

Delhi Public School Kalyanpur, Kanpur — Uttar Pradesh, India
- *Intermediate: 95.6%* — *2019*
  High School: 10 CGPA — *2017*

SKILLS SUMMARY

- **Languages**:   Python, C/C++, Java, SQL, Bash
- **Frameworks**:   PyTorch, TensorFlow, Keras, ScikitLearn, Transformers, OpenCV, Django, React, Pandas
- **Tools & Tech**: Docker, GIT, LLMs, LaTeX, Linux, MATLAB
- **Coursework**:   Artificial Intelligence, Operating Systems, Data Analytics & Visualisation, Data Structures, Algorithms, DBMS (SQL), Cryptography, Machine Learning, Probability, Linear Algebra

EXPERIENCE

- **Ripik.AI** — Noida, Uttar Pradesh
  Data Scientist — *November 2023 - Present*
  – Developing a **Generative AI** application focused on implementing a **RAG system** for PDF document ingestion and retrieval. Leveraging **GPT-3.5** as the language model, **Pinecone** for Vector Storage, and **RAPTOR** for efficient retrieval. Actively engaging in testing diverse technologies to enhance current system performance.
  – Previously Built and Optimized ML models for tasks such as **Parameter Optimization**, **SOPs Generation**, etc. using Tensorflow, scikit-learn, etc. and helped deliver end-to-end solutions.

- **Arista Networks** — Remote
  Software Engineering Intern — *May 2022 - July 2022*
  Research, Development & Testing of new **Wireless Intrusion Prevention System** (WIPS) mechanisms as an intern. The aim was to sever unauthorized connections between **Access Points(APs) and Clients** in a Network since the methods that were used with the 802.11 standard would not work after the **802.11w** amendment.

KEY PROJECTS & RESEARCH

- **Machine Learning for Improving Document-Reranking** — *IIT Bhilai*
  AI Research under Dr. Soumajit Pramanik (IIT Bhilai) & his Colleagues from France — *2023*
  Worked on a system to re-rank documents using a **combination** of Knowledge Graph along with textual data to give better and more relevant search results. Used a combination of **Machine Learning (NLP) & Information Retrieval** techniques involving GCNs, LSTMs, BERT Embeddings, etc.

- **AI Sudoku-Solver** — *IIT Bhilai*
  Deep Learning & Computer Vision — *2022*
  Developed a **Python** application to detect Sudoku from an image, and output a solution using Deep Learning & Computer Vision technologies. **Tensorflow** and **Keras** were used to generate a model based on **Neural Networks** for the recognition of Digits from Sudoku. Used **OpenCV** library to read images and implemented **Contour Detection** to identify Sudoku from the image.

- **KD & PEFT on BERT** — *IIT Bhilai*
  LLM Fine Tuning & Compression — *2023*
  Utilized PEFT (LoRA) to **fine-tune** BERT for the **Question Answering** task, leveraging the HuggingFace transformers library. Additionally, employed **Knowledge Distillation** to compress the model, resulting in a Student Model **10x smaller** in size while having comparable accuracy, as tested on the SQuAD dataset.

MISCELLANEOUS

- **Codeforces** Personal Highest Rating **1463** (**Specialist**)
  Global Rank: **791** in Round #760
- **Codechef** Personal Highest Rating **1767**
  Global Rank: **83** in Codechef May Cook-Off 2021 Division 3
- **Position of Responsibility**: Executive at National Service Scheme, IIT Bhilai
- Ranked among the top **4%** of all who were eligible for JEE Advanced 2019 and
  Among the top **4 %** in JEE Mains 2019