Akshit Kumar | CV

IIIT Hyderabad – Telangana State – India

☐ +91 9622690011 • ✉ akshit.kumar@research.iiit.ac.in • ○ komikat

Education

International Institute of Information Technology, Hyderabad   2021–
B.Tech in Computer Science, MS in Computational Linguistics (by Research)   *Hyderabad, IN*
7.17/10

Delhi Public School, RK Puram   April 2021
All India Senior School Certificate Examination   *New Delhi, IN*
96/100

Experience

Research

Language Technologies Research Center   **IIIT Hyderabad**
Undergraduate Student Researcher, under the guidance of Prof Parameshwari Krishnamurthy   *2023–*
– I work on parsing natural language (with an emphasis on Indian languages).
– Trained and presented findings on an array of SOTA models including XLM Roberta and self-attention models written from scratch.
– Wrote a multilingual dependency parser for parsing paninian treebank of Indian languages.

Open Source

Haskell.org   *Haskell language server*
Google Summer of Code participant   *2024*
– HLS is a server implementation of the Language Server Protocol for the Haskell programming language.
– The server "talks" to a client (a text editor for instance) which can ask the server to do operations like reporting errors, providing code completions.
– Overhauled the HLS test suite – making the tests faster, less flakier and parallel execution ready.

Projects

Grounding small LMs on structured data using Multi hop Question Generation   **Information Retrieval course project**
Pytorch, chromadb   *2024*
A novel pipeline to extract reasoning chains from knowledge graphs and use a teacher LM to guide small LM to reason.

Multidomain Black-Box Machine-Generated Text Detection   **Advanced NLP Project**
PyTorch, pandas   *2023*
Developed a robust system for the detection of machine-generated text by comparing multiple models: baseline transformer, fine-tuned RoBERTa and LoRa on RoBERTa.

Codemixed text generation   **Intro NLP Project**
PyTorch, pandas   *2023*
Compared different techniques to generate a Hindi-English (Hinglish) code-mixed sentence given a sentence in English: baseline LSTM, IndicBART, mT5.

RL based UAV controller   **RL course project**
PyTorch, matplotlib   *2024*
Used **deep Q-learning** to predict optimal control input for a quadcopter UAV.

Political Bias Analysis   **CL course project**
sklearn, PyTorch, spacy   *2022*
Trained multiple models (Multinomial and Bernoulli Naive Bayes, RNN, tf-idf) on annotated newpaper articles and twitter threads to analyse political bias in media.

Compiler   **Compilers course project**
Racket   *2024*
A compiler in racket for a strictly typed language; written in multiple nanopasses including **register allocation** using an interference graph.

Network File system   **OS course project**
C, Make   *2023*
A network file system from scratch in C; supports **parallel clients** and storage servers; uses an **LRU cache** for efficient queries.

Greddiit – full stack web project   **Software systems course project**
ReactJS, ExpressJS, MongoDB, Docker   *2022*
A fully featured reddit clone including profiles, admin controls and infinitely nested comments support

Skills & Interests

ML: PyTorch, pandas, NumPy, sklearn, spacy

Dev: React, React Native, AWS, Docker, Typescript, ExpressJS, Django, Flask, ZSH, MongoDB, PostgreSQL, Git, Jupyter, VS Code

Misc: Emacs, Linux, Maths, Competitive programming, Functional Programming, Russian Literature, Garage rock, Hiking