

Tanmay Sachan

Mail: tanmay.sachan@research.iiit.ac.in
Phone No.: +91 8700380515
LinkedIn:// [tanmay-sachan](#)
Github:// [tanmaysachan](#)
Website: [tanmaysachan.github.io](#)

EDUCATION

IIIT HYDERABAD

B.TECH AND M.S. IN COMPUTER
SCIENCE AND ENGINEERING
Hyderabad, Telangana
CGPA: 9.03/10.0

COURSEWORK

Data Structures and Algorithms
Software Programming for Performance
Operating Systems
Databases
Machine, Data and Learning
Natural Language Processing

SKILLS

Over 10000 lines:

• Python • C++ • C

Over 1000 lines:

Rust • Bash • Go • Erlang

JavaScript

Familiar:

Haskell • Julia • Ruby • SQL

Libraries, Frameworks:

Pytorch • TensorFlow • Django • NLTK

OpenMP • CUDA • ReactJS

ACHIEVEMENTS

ACM-ICPC Regionals - 2018

Selected for the ACM-ICPC regionals,
ranked 128 out of over 2000 participating
teams.

IOI-TC - 2017,2018

Selected twice for the International
Olympiad in Informatics(IOI) training
camp, among top 30 school programmers
in India.

POSITIONS HELD

Open Source Development Group
(Admin), IIIT-H

Tasked with holding sessions on various
topics relating to Open source and
Software development.

Teaching Assistant, IIIT-H

Leading a research project on political
leaning in text documents for a group of
students.

PUBLICATIONS

September-2021

- "PreCogIIITH at HinglishEval: Leveraging Code-Mixing Metrics Language Model Embeddings To Estimate Code-Mix Quality" published at INLG 2022.
- "SCATE: Shared Cross Attention Transformer Encoders for Multi-Modal Fake News Detection" published at ASONAM 2021.

RESEARCH

GEORGIA TECH & PRECOG, IIIT-D

August-2022 - Present

Recognizing dark behaviour of applications on Play store

- Working on developing **ML fingerprinting** techniques for classifying application behaviours as potentially fraudulent.

PRECOG, IIIT-D

December-2021 - March-2022

Fraud detection through Identity Resolution

- Working with internal teams of e-commerce companies on creating **identity resolution graphs** across social media platforms, to enable better detection of **fraudulent users** on e-commerce websites.

INFORMATION RETRIEVAL AND EXTRACTION LAB, IIIT-H

May-2020 - Present

Multi-Modal summarization (Publication under review at EMNLP 2022)

- Worked on improving **cross-modal** correspondence between images and text within news articles through **semantic multimodal content reranking** and **knowledge distillation**.

EXPERIENCE

APPLE | APPLIED RESEARCH ENGINEER

August-2022 - Present

- Designing **scalable algorithms** to cluster and derive business insights from geospatial data.
- Building interpretable **ML** models to score actionability of reported issues.
- Using **PyTorch**, **Scikit**, **Flask** and **Docker** for training and managing model deployments.

LINKEDIN | SOFTWARE ENGINEER INTERN (MACHINE LEARNING)

May-2022 - August-2022

- Worked in the Bangalore AI team to generate dense user activity representations for **improving content ranking** for LinkedIn-learning search.
- Modeled and trained state of the art **Deep Learning architectures** for embedding time series textual data corresponding to user activity.
- Project ramped up to production.

SUBTL.AI | NLP ENGINEER

October-2021 - January-2022

- Worked on improving **semantic search** over unstructured bank documents through state of the art **NLP** models and techniques.
- Made use of **PyTorch** for model creation and training, and **Docker**, **Kubernetes** for scaling the models to production.

GOOGLE SUMMER OF CODE | SOFTWARE DEVELOPER | MENTOR

May-2019 - August-2019

- Worked on **parallelising** the test suite of the Xapian Search Engine Library, **improving runtime by ~20%**.
- Improved I/O profiling by implementing **system call tracing**.