

TANMAY SACHAN

@ tnmysachan@gmail.com +91 8700 380 515 tanmaysachan.github.io

PUBLICATIONS

- "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 EXPERIENCE

Precog

Jan 2022 – Present Hyderabad, India

- Collaborated with Georgia-Tech on developing **ML fingerprinting** techniques for classifying application behaviours as potentially **fraudulent** (work under review at PAKDD 2023).
- Developed techniques for **qualitative evaluation of code-mixed English-Hindi data**.
- Worked 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.

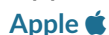
Information Retrieval and Extraction Lab, IIIT-H

Apr 2020 – Present Hyderabad, India

- Building a large **Multimodal + Multilingual Parallel corpus** for 12 different Indian Languages.
- Improved **cross-modal** correspondence between images and text through **semantic content reranking** and **knowledge distillation** for the task of **Multimodal Summarization** (work under review at EACL 2023).
- Boosted the performance of Multimodal fake news detection models. Proposed an architecture that **raised accuracy by 3%** across several datasets.

WORK EXPERIENCE

Applied Research Engineer Intern



Aug 2022 – Present Hyderabad, India

- Designed **scalable algorithms** to cluster and derive business insights from geospatial data.
- Constructing **interpretable ML models** to score actionability of reported issues on **Apple Maps**.
- Enhanced clustering of priority **Multilingual** issues, improving their visibility **by over 10%**.

Machine Learning Intern



May 2022 – Aug 2022 Bangalore, India

- Worked in the AI team on improving user representations for **content ranking** on LinkedIn-learning.
- Modeled and trained state of the art **Deep Learning architectures** for embedding time series textual data corresponding to user activity.
- Raised search metrics - **Mean Reciprocal rank by 4.4%** and **Normalized Discounted Cumulative Gain by 2.3%**, project ramped to production.

EDUCATION

B.Tech. and M.S. in Computer Science and Engineering (Integrated)

International Institute of Information Technology, Hyderabad

2018 – 2023 Hyderabad, India

CGPA - 9.03/10

RELEVANT COURSEWORK

Optimization Methods
Natural Language Processing
Information Retrieval and Extraction
Statistical Methods in AI
Machines, Data and Learning

ACHIEVEMENTS

2x Merit List, 1x Dean's List
Awarded for outstanding academic performance.

ACM-ICPC Regionals - 2018
Qualified by ranking 40 out of over 3000 participating teams.

IOI-TC - 2017, 2018
Qualified twice for the International Olympiad in Informatics training camp, part of top 30 highschool level programmers in India.

SKILLS

C++ Python PyTorch Tensorflow
Machine Learning Multimodality
Communication Leadership

POSITIONS HELD

Open Source Development Group (Admin)

Aug 2020 – Aug 2022

- Held informational sessions and hackathons on topics relating to Open source and Software development.

Teaching Assistant

Aug 2021 – Dec 2021

- Managed a class of 100 for the course **Information Retrieval and Extraction**.
- Mentored a project on detection of political bias in text.