

Punyajoy Saha

Email: punyajoy@iitkgp.ac.in

Online presence:
[Linked in](#), [Researchgate](#), [Github](#)
Website: <https://punyajoy.github.io/>

Research Experiences

MAY 2019 - DEC 2019

- **CNeRg Lab, IIT Kharagpur**– JRF sponsored by the project “Cognitive Stimuli: Detecting and Generating Exaggeration in Online and Social Media Content” (CSAM).

AUGUST 2019 - SEPTEMBER 2019

- **Language Technology Group, University of Hamburg, Germany**–Intern, worked on developing the backend for Forum 4.0 project.

MAY 2018 - APRIL 2019

- **CNeRg Lab, IIT Kharagpur**–Intern, worked on Detection of Hate Speech in Online Social Networks.

MAY 2017 - JULY 2017

- **Neurocomputing Lab, IIT Delhi** –Intern, worked on Analysis of Hindi letters and Hindi documents comprehension using Eye Tracking.

Education

JAN 2020 - Present

Indian Institute of Technology, Kharagpur – 2nd year PhD student, Computer Science and Engineering, IIT Kharagpur.

- Current CGPA - 9.13
- Member of CNeRg lab.
- Working on analysis, detection and countering hatespeech and similar toxic language in social media.

JULY 2015 - MAY 2019

Indian Institute of Engineering Science and Technology, Shibpur – B.Tech in Computer Science and Technology

- Graduated with CGPA of 8.93
- Computer Vision Head at Robodarshan, the robotics club of Indian Institute of Engineering, Science and Technology Shibpur(2017-18)
- BTech Project - Application of Generative Adversarial Networks for Wallpaper Generation on Digital Devices ([Link](#))

Awards and Achievements

- Won Ted Nelson newcomer award for “You too Brutus! Trapping Hateful Users in Social Media: Challenges, Solutions & Insights” at ACM Hypertext 2021.

- 1st in both subtasks of Abusive and Threatening Language Detection for Tweets in Urdu: Abusive Language Detection at FIRE 2021. [Link](#)
- Received [PMRF scholarship](#) for supporting my PhD.
- 1st, 2nd and 1st Rank in different languages in Offensive shared task at [DravidianLangTech-2021](#) ([preprint](#), [code](#))
- World Rank 11 in Facebook Hatememe detection competition ([link](#))
- Winner of the worldwide Facebook+Social Science One data [grant](#).
- Winner of offensive speech detection competition for the German language at [HASOC 2019](#).
- The preprint "Hateminers: Detecting hatespeech against women" got recognised as the best preprint of the week. [See here](#)
- Winner of the misogynistic text classification competition for the English language in [AMI Evalita-2018](#). ([preprint](#), [code](#))

Publications

1. The paper "You too Brutus! Trapping Hateful Users in Social Media: Challenges, Solutions & Insights" was accepted at [ACM hypertext 2021](#).
2. The paper "Hate-Alert@DravidianLangTech-EACL2021: Ensembling strategies for Transformer-based Offensive language Detection" was accepted at First [Workshop](#) on Speech and Language Technologies for Dravidian Languages, EACL 2021. ([paper](#), [code](#))
3. The paper "Short is the Road that Leads from Fear to Hate": Fear Speech in Indian WhatsApp Groups" was accepted at [The Web Conference-2021](#) (erstwhile WWW) ([paper](#), [code](#))
4. The paper "HateXplain: A Benchmark Dataset for Explainable Hate Speech Detection" was accepted at [AAA1-2021](#) ([paper](#), [code](#))
5. The paper "Using Knowledge Graphs to Improve Hate Speech Detection" was accepted at [CODS COMAD YSR 2021](#)
6. The paper "Hate begets Hate: A Temporal Study of Hate speech" was accepted at [CSCW 2020](#) ([paper](#), [code](#))
7. The paper "Deep learning models for multilingual hate speech detection" was accepted at [ECML-PKDD 2020](#) ([paper](#), [code](#))
8. The paper "HateMonitors: Language Agnostic AbuseDetection in Social Media" was accepted at HASOC track in [FIRE 2019](#), Kolkata. ([paper](#), [code](#))
9. The paper "Thou shalt not hate: Countering Online Hatespeech" was accepted at the international conference- [ICWSM-2019](#), Munich, Germany. ([paper](#), [code](#)).
10. The paper "Effect of Devanagari Font Type in Reading Comprehension: An Eye-Tracking Study" was presented at the international conference [IHCI-2018](#), Allahabad, India. The paper was also a part of the book "Lecture Notes in Computer Science" published by Springer. [Paper](#).
11. The paper "Document categorisation using graph structuring" was presented at the international conference [ICCACP-2017](#), Sikkim, India. [Paper](#)

Talks and Tutorials

- Invited talk on the topic of “Generative Adversarial Network” IEST, Shibpur, 2020.[Link](#)
- Tutorial on “Hate speech: Detection, Mitigation and Beyond” accepted at ICWSM 2021. [Link](#)
- The paper “Short is the Road that Leads from Fear to Hate”: Fear Speech in Indian WhatsApp Groups” will be presented at [IC2S2](#) ([preprint](#), [code](#))
- Tutorial on “Hate speech: Detection, Mitigation and Beyond” accepted at AAAI 2022. [Link](#)
- Invited talk in Understanding and Automating Counterspeech workshop, 2021 organised by CRASSH

Referees

- Prof. Animesh Mukherjee, Associate Professor in Computer Science and Engineering, IIT Kharagpur, animeshm@cse.iitkgp.ac.in.
- Prof. Pawan Goyal, Assistant Professor in Computer Science and Engineering, IIT Kharagpur, pawang.iitk@gmail.com.
- Prof. Chris Biemann, Professor for Language Technology, University of Hamburg, biemann@informatik.uni-hamburg.de
- Kiran Garimella, Michael Hammer Fellow, MIT Institute for Data, Systems, and Society, garimell@mit.edu
- Prof. Scott Hale, Associate professor at the Oxford Internet Institute, scott.hale@oii.ox.ac.uk