

Aswini Kumar Padhi

in: <https://www.linkedin.com/in/aswini-kumar-08925572/>

Q: <https://github.com/AswiniNLP>

Link for Personal Website

Email : aswinikumarpadhi1995@gmail.com

Mobile : +91-7978014958

EDUCATION

- **Indian Institute of Technology, Delhi** New Delhi, India
Doctor of Philosophy in Natural Language Processing; GPA: 9.00/10 Jul 2023 - Present
 - *Courses Credited: Social Network Analysis (ELL880), Data Visualization (ELL824), Deep Learning for Natural Language Processing (ELL884), Cloud Computing (ELL887)*
- **ABV-Indian Institute of Information Technology, Gwalior** Gwalior, India
Master of Technology in Computer Science and Engineering; GPA: 9.34/10 Aug 2021 - May 2023
 - *M.Tech Thesis: "I2C SLR: Indian Continuous Sign Language Recognition System"*
 - *Ranked 1st in Department on the basis of GPA*
- **Veer Surendra Sai University of Technology, Burla** Burla, India
Bachlore of Technology in Computer Science and Engineering; GPA: 8.4/10 Aug 2012 - May 2016
 - *B.Tech Thesis: "Design of 4-Bit Microprocessor"*

SKILLS SUMMARY

- **Languages:** Python, C++, C, Java, SQL, Unix scripting, Vega-Lite
- **Libraries & Frameworks:** PyTorch, PyTorch Geometric, TensorFlow, Keras, Scikit-Learn, Numpy, Pandas, OpenCV, nltk
- **Tools:** Git

PUBLICATIONS

*: Equal contribution

- [C1] **Aswini Padhi***, Anil Bandhakavi, Tanmoy Chakraborty, "Counterspeech the ultimate shield! Multi-Conditioned Counterspeech Generation through Attributed Prefix Learning", **ACL 2025 (Main)**, link May 2025 (**Core A* conference**).
- [C2] Amey Hengle*, **Aswini Padhi***, Anil Bandhakavi, Tanmoy Chakraborty, "CSEval: Towards Automated, Multi-Dimensional, and Reference-Free Counterspeech Evaluation using Auto-Calibrated LLMs", **NAACL 2025 (Main)**, link February 2025.
- [C3] Amey Hengle*, **Aswini Padhi***, Sahajpreet Singh, Anil Bandhakavi, Md Shad Akhtar, Tanmoy Chakraborty, "Intent-conditioned and Non-toxic Counterspeech Generation using Multi-Task Instruction Tuning with RLAIIF", **NAACL 2024 (Main)**, link, March 2024.
- [C4] Subodh Rajpopat, **Aswini Kumar Padhi**, Sunil Kumar, "An Efficient Correlated Gamma Wave-Based Lightweight Model for Schizophrenia Detection", **CVMI 2023 (Main)**, link, December 2023.

RESEARCH PROJECTS

- **Multi-Conditioned Counterspeech Generation through Hierarchical Prefix Learning** IIT, Delhi
Ph.D. Research work; Supervisor - Dr. Tanmoy Chakraborty (Laboratory of Computational Social Systems) May 2023 - May 2027
 - This work titled "Counterspeech the Ultimate Shield! Multi-Conditioned Counterspeech Generation through Attributed Prefix Learning" was accepted to ACL 2025, the 63rd Annual Meeting of the Association for Computational Linguistics, in the main conference.
 - In this work, we proposed a two-phase framework named **HiPPrO** that generates attribute-guided counterspeech by hierarchically learning prefix embeddings for intent and emotion, followed by reference- and reward-free preference optimization for response alignment.
 - We curated **MultiCONAN**, an enhanced version of the IntentCONANv2 dataset, which includes 13,973 counterspeech instances annotated with both intent and emotion labels by expert annotators, supporting fine-grained generation control.
 - Extensive evaluations show that HiPPrO significantly outperforms baseline and state-of-the-art models (e.g., CoARL, GPT-3.5, GPT-4), achieving up to **38% improvement in Intent Conformity** and outperforming baselines across ROUGE and BERTScore metrics, while maintaining low toxicity and strong human evaluation rankings.

- This work contributes to the development of safer and more context-aware language agents, offering a scalable approach for generating respectful and persuasive responses in sensitive social scenarios.
- **Exposure:** PyTorch, HuggingFace, PEFT, FLAN-T5, Preference Optimization, Chain-of-Thought Prompting.

• **Effective Counter-Speech Evaluation using LLMs**

IIT, Delhi

Ph.D. Research work; Supervisor - Dr. Tanmoy Chakraborty (Laboratory of Computational Social Systems) May 2023 - May 2027

- This work titled "CSEval: Towards Automated, Multi-Dimensional, and Reference-Free Counterspeech Evaluation using Auto-Calibrated LLMs" was accepted to NAACL 2025, the North American Chapter of the ACL (Association for Computational Linguistics), in the main conference.
- In this work, we have developed a novel framework called Auto-CSEval for evaluating the quality of automated counterspeech generation across four key dimensions: contextual relevance, aggressiveness, argument coherence, and suitability. Auto-CSEval uses auto-calibrated chain-of-thoughts (CoT) to improve alignment with human judgment.
- We have curated a large benchmark dataset named CSEval, containing 7,926 model-generated counterspeeches with expert human assessments across the four quality dimensions. The dataset includes 2,223 unique hate speech instances and 4,318 ground-truth (reference) counterspeeches.
- We found that our Auto-CSEval method consistently outperformed existing evaluation metrics, showing significant improvements in correlation with human judgment. Specifically, it demonstrated an average improvement of 4.8 percentage points in Spearman correlation and 4.9 percentage points in Kendall-Tau correlation compared to the best existing methods.
- This work contributes to research in the domain of Artificial Intelligence for Social Good, specifically addressing the challenge of evaluating automated counterspeech generation to combat online hate speech.
- **Exposure:** PyTorch, HuggingFace, Large Language Models.

• **Effective Counter-Speech Generation Using Reinforcement Learning**

IIT, Delhi

Ph.D. Research work; Supervisor - Dr. Tanmoy Chakraborty (Laboratory of Computational Social Systems) May 2023 - May 2027

- This work titled "Intent-conditioned and Non-toxic Counterspeech Generation using Multi-Task Instruction Tuning with RLAI¹", was accepted in NAACL¹, North American Chapter of the ACL (Association for Computational Linguistics) 2024, in the main conference.
- In the work we have designed a novel model for generating effective and non-toxic counter speeches for a hate speech conditioned on four different intents. Reinforcement learning with AI feedback (RLAIF) made our model generate more non-toxic counter-speeches.
- We have curated a huge hate speech counter-speech pair dataset named IntentCONAN V2. In this dataset, there are ~4000 unique hate speech instances and each hate speech has four distinct intents accordingly we have ~16000 counter-speeches instances. All our annotators have gone through proper counter-speech generation documentation before starting the data annotation task.
- We found that our model performed especially well in intent conformity ~3% and argument quality ~4% on generated counter speeches in the test set.
- This work involved research work in the domain of Artificial intelligence for Social Good.
- **Exposure:** PyTorch, HuggingFace, TRL, Large Language Models.

• **I2C SLR: Indian Continuous Sign Language Recognition**

IIITM, Gwalior

M.Tech. Thesis; Supervisor - Dr. Sunil Kumar (Computer Vision and Multimedia Laboratory)

Jan 2022 - May 2023

- This work titled "A Pose-based Transfer learning Approach with Connectionist Temporal Loss for Indian Continuous Sign Language Recognition System", is under review phase.
- We designed a novel sign language recognition model that effectively produces the sign glosses from a continuous sign language video. The model first trained on recognizing the Indian isolated sign languages and then the learning method incorporated into finding the sign boundaries in a continuous sign language video.
- We have evaluated our approach on the popular open-sourced ISL-CSLTR: Indian Sign Language Dataset for Continuous Sign Language Translation and Recognition dataset.
- We found that our model performed especially well in recognition of sign glosses achieving the lowest possible word error rate of ~9% in the test set.
- This work involved research work in the domain of the Indian sign language community.

¹<https://arxiv.org/abs/2403.10088>

- **Exposure:** PyTorch, OpenCV, Large Language Models, Recurrent Neural Networks, OpenHands.

• An Efficient Lightweight Model for Schizophrenia Detection

IITM, Gwalior

Research Work - Dr. Sunil Kumar (Computer Vision and Multimedia Laboratory)

August 2021 - June 2022

- This work titled “An Efficient Correlate Gamma wave-based lightweight model for Schizophrenia detection”, was published in the IEEE International Conference on Computer Vision and Machine Intelligence ².
- In this work we designed a novel lightweight Schizophrenia detection model with the help of Gamma wave extraction from the electroencephalogram brain waves.
- The proposed model has approximately 12500 parameters with an accuracy of ~99.5% which is much lighter than the previous state-of-the-art models that contain nearly 11 million parameters.
- This work involved research work in the healthcare area.
- **Exposure:** Pytorch, OpenCV

SERVICES

- Teaching Assistant for *CEP-GenAI* course - 2025-26 at *IIT Delhi*
- Teaching Assistant for *ELL884 (Deep Learning for NLP)* course in Sem-4 of 2025-26 at *IIT Delhi*
- Teaching Assistant for *ELL409 (Machine Learning)* course in Sem-3 of 2024-25 at *IIT Delhi*
- Teaching Assistant for *ELL101 (Introduction to Electrical Engineering)* course in Sem-I of 2023-24 at *IIT Delhi*

WORK EXPERIENCE

• Wipro Technologies

Bengaluru, India

Associate Consultant on Machine Learning for Client: Chevron USA

December 2016 - June 2019

- **Job Role: Machine Learning-Based Predictive Maintenance System for Oil Equipment:** Remotely led the development of a machine learning model for Chevron's oil equipment monitoring from the offshore delivery center in India. Collaborated virtually with on-site engineers to understand the data collected from IoT sensors on approximately 1,000 pieces of rotating equipment. Developed and fine-tuned a random forest algorithm using Python and scikit-learn to predict potential equipment failures based on historical sensor data. The model analyzed parameters such as vibration patterns, temperature fluctuations, and pressure readings to identify early signs of malfunction.

• Wipro Technologies

Hyderabad, India

Associate Consultant on Machine Learning for Client: VALE S.A.

July 2019 - December 2020

- **Job Role: AI-Powered Inventory Optimization for Iron Ore Supply Chain:** Managed the creation of an AI-powered inventory optimization system for VALE SA's iron ore supply chain from the offshore center in India. Worked remotely with the client's logistics team to gather historical inventory and shipment data. Developed a deep learning model using TensorFlow to forecast iron ore demand across different global markets. The model incorporated external factors such as market prices, seasonal trends, and macroeconomic indicators to improve prediction accuracy. Implemented the model within VALE's existing ERP system through API integrations, providing real-time inventory recommendations.

HONORS AND AWARDS

- Received Microsoft Research India PhD Fellowship 2025 by Microsoft Research India (MSRI).
- Selected for Microsoft Research India travel grant to travel to Mexico City for NAACL 2024.
- Ranked 1st in the Department of Computer Science and Engineering, IITM (Gwalior), among the post-graduating batch of 2023
- Qualified among the top 2% of the students (about 160,000) appearing for *Graduate Aptitude Test in Engineering (GATE)*, 2021
- Received National Cadet Corps (NCC) 'C' certificate during *NCC Examination*, 2010

²https://scholar.google.com/citations?view_op=view_citation&hl=en&user=SNF6BsQAAAAJ&citation_for_view=SNF6BsQAAAAJ:2osOgNQ5qMEC