Kaushal Kumar Maurya

@ kaushalmaurya.iith@gmail.com   🌐 https ://kaushal0494.github.io/   📱 +(91) 94410 58374   Github   🎓 Google Scholar
📍 Unit-515, Leonardo Residence, Block-B, Masdar City, Abu Dhabi, UAE   in Linkedin   🐦 Twitter

🔍 Research Interests

Multilingual NLP, Responsible NLP, Educational NLP, Natural Language Processing, and Machine Learning

🎓 Education

2023	Ph.D. in Computer Science and Engineering, Indian Institute of Technology Hyderabad (IITH), India
2017	Masters in Artificial Intelligence, SCIS, University of Hyderabad, India
2014	Bachelors in Computer Science and Engineering, KNMIET, UPTU, India

📖 Publications

Theses
> [Ph.D. Thesis] Extending Generative NLP : Incorporating Diversity, Context, and Inclusivity in Neural Text Generation
> [Masters Thesis] The Study of Machine Translation Evaluation

Conferences
> **Kaushal Kumar Maurya**, KV Aditya Srivatsa, Kseniia Petukhova, and Ekaterina Kochmar. "Unifying AI Tutor Evaluation : An Evaluation Taxonomy for Pedagogical Ability Assessment of LLM-Powered AI Tutors" at *NAACL 2025* 🔗[Paper]
> **Kaushal Kumar Maurya**, KV Aditya Srivatsa, and Ekaterina Kochmar. "SelectLLM : Query-Aware Efficient Selection Algorithm for Large Language Models." Under Review. 🔗[Paper]
> Aishwarya Maheswaran, **Kaushal Kumar Maurya**, Manish Gupta, and Maunendra Sankar Desarkar. "DAC : Quantized Optimal Transport Reward-based Reinforcement Learning Approach to Detoxify Query Auto-Completion" at A*CM SIGIR 2024*. 🔗[Paper]
> Aishwarya Maheswaran*, **Kaushal Kumar Maurya***, Manish Gupta, and Maunendra Sankar Desarkar. "DQAC : Detoxifying Query Auto-completion with Adapters" at *PAKDD 2024.* 🔗[Paper]
> **Kumar Kumar Maurya**, Rahul Kejriwal, Maunendra Desarkar, and Anoop Kunchukuttan. "CharSpan : Utilizing Lexical Similarity to Enable Zero-Shot Machine Translation for Extremely Low-Resource Languages" at *EACL 2024*. 🔗[Paper]
> Maharaj Brahma*, **Kaushal Kumar Maurya***, and Maunendra Sankar Desarkar. "SelectNoise : Unsupervised Noise Injection to Enable Zero-Shot Machine Translation for Extremely Low-Resource Languages" at *Findings of EMNLP 2023.* 🔗[Paper]
> **Kumar Kumar Maurya**, Maunendra Sankar Desarkar, Manish Gupta and Puneet Agrawal, "Trie-NLG : Trie Context Augmentation to Improve Personalized Query Auto-Completion for Short and Unseen Prefixes" at *ECML-PKDD 2023 (Journal Track : DAMI)*. 🔗[Paper].
> Venkatesh E, **Kaushal Kumar Maurya**, Deepak Kumar, Maunendra Sankar Desarkar, "DIVHSK : Diverse Headline Generation using Self-Attention based Keyword Selection" at *Findings of ACL 2023*. 🔗[Paper].
> **Kaushal Kumar Maurya**, Maunendra Sankar Desarkar, "Meta-XNLG : A Meta-Learning Approach Based on Language Clustering for Zero-Shot Cross-Lingual Transfer and Generation" at *Findings of ACL 2022*. 🔗[Paper].
> Kamal Shrestha, Aditi Sanjay Bagora, **Kaushal Kumar Maurya**, Maunendra Sankar Desarkar, "Hostility Detection in Online Hindi-English code-mixed Conversations" at *WebSci 2022*. 🔗[Paper].
> **Kaushal Kumar Maurya**, Maunendra Sankar Desarkar, Yoshinobu Kano, Kumari Deepshikha, "ZmBART : An Unsupervised Cross-lingual Transfer Framework for Language Generation" at *Findings of ACL 2021*. 🔗[Paper].
> **Kaushal Kumar Maurya**, Maunendra Sankar Desarkar, "Learning to Distract : A Hierarchical Multi-Decoder Network for Automated Generation of Long Distractors for Multiple-Choice Questions for Reading Comprehension" at *CIKM 2020*. 🔗[Paper].
> **Kaushal Kumar Maurya**, Renjith P. Ravindran, Ch Ram Anirudh, Kavi Narayana Murthy, "Machine Translation Evaluation : Manual Versus Automatic—A Comparative Study" at *DECT 2020*. 🔗[Paper].

Journals	> Sreekanth Madisetty, **Kaushal Kumar Maurya**, Akiko Aizawa, Maunendra Sankar Desarkar "A Neural Approach for Detecting Inline Mathematical Expressions from Scientific Documents" at *Expert Systems 2020*. [Paper].
Workshop	> KV Aditya Srivatsa, **Kaushal Kumar Maurya**, and Ekaterina Kochmar. "Harnessing the Power of Multiple Minds : Lessons Learned from LLM Routing" at Workshop on Insights from Negative Results, *NAACL 2024.* [Paper]
	> **Kaushal Kumar Maurya** and Maunendra Desarkar. "Towards Low-resource Language Generation with Limited Supervision" at *Big Picture Workshop, EMNLP 2023*. [Paper]
	> Arkadipta De, Venkatesh E, **Kaushal Kumar Maurya**, and Maunendra Sankar Desarkar, "Coarse and Fine-Grained Hostility Detection in Hindi Posts using Fine Tuned Multilingual Embeddings" at *CONSTRAINT workshop, AAAI 2021*. [Paper]. Shared task best paper honorable mention.
Tutorials	> Bashar Alhafni, Sowmya Vajjala, Stefano Bannò, **Kaushal Kumar Maurya**, and Ekaterina Kochmar. "LLMs in Education : Novel perspectives, challenges, and opportunities." at *COLING 2025* [Tutorial Webpage].
System Demonstrations	> **Kaushal Kumar Maurya** and Ekaterina Kochmar "`AITutor-AssessmentKit` : An Open-Source Library to Measure Pedagogical Ability of AI Tutors in Educational Dialogue." Under Review.

💼 PROFESSIONAL ENGAGEMENTS

Experience	**Post-Doctoral Research Associate, MBZUAI, Abu Dhabi, UAE, JANUARY 2024 - PRESENT** PI : Dr. Ekaterina Kochmar > *Research Theme :* AI/LLM for Education > Broadly engaged in the development of an LLM-powered co-pilot for education, with a specific focus on creating intelligent tutoring systems, developing comprehensive evaluation frameworks, and establishing benchmarks to track and accelerate progress in Educational NLP. **Data Scientist, Ntwist, CIE, IIITH, Hyderabad, India, JULY 2017 - MARCH 2018** > Optimization and automation of gas-plant production by leveraging deep reinforcement learning. Conducted correlation study of attributes in higher dimensions.
Internships	**Microsoft India (R&D), AUGUST 2022 - DECEMBER 2022** Topic : Machine Translation for Extremely Low-Resource Languages/Dialects > Proposed a novel MT system for extremely low-resource languages (LRLs) and dialects by exploiting lexical similarity with closely related high-resource languages. > **Publication :** EACL 2024 **Microsoft India (R&D), JUNE 2021 - JULY 2021** Topic : Personalized Auto-suggest Text Generation for Bing Search Engine > Augmentation of trie suggestions/completions in an NLG model for personalized auto-suggestion tasks for Bing Search Engine. > **Publication :** ECML-PKDD 2023 **Nvidia AI Research Center, India, and Shizuoka University, Japan, MAY 2020 - JANUARY 2021** Topic : Towards Cross-lingual Transfer and Generation for Indian Languages > Proposed an unsupervised cross-lingual generation framework, ZmBART, to transfer supervision from a high-resource language to multiple low-resource languages. > **Publication :** Findings of ACL 2021
Collaboration	**Microsoft India (R&D), Academic Grant, JULY 2021 - JUNE 2023** Topic : Non-toxic Multilingual Personalized Auto-suggest Generation > Improved the quality of auto-suggestions for Bing Search Engine. > Overcame problematic/toxic auto-suggestions for Bing Search Engine. > Enabled auto-suggest generation in low-resource languages with limited supervision. > **Publication :** SIGIR 2024 and PAKDD 2024

Startup Advisor	Technical Advisor at Lending Katalyst [AI-NLP Startup in Legal Domain], FEBRUARY 2022 - JUNE 2023
	Topic : Non-toxic Multilingual Personalized Auto-suggest Generation
	› Provided AI/NLP-oriented solutions for real-world use cases in the legal domain.
	› Addressed interdisciplinary problems involving NLP, multilingualism, the legal domain, and linguistics.
Podcast	With Priya , OCTOBER 2023
	Topic : AI Research and My PhD Research Experience
	› Link : ▶ YouTube
Invited Talks	NLP Team, National Research Council, Canada, OCTOBER 2024
	› Title : Towards Machine Translation for Extremely Low Resource Languages
	NLP Seminar Series, MBZUAI, Abu Dhabi, UAE, AUGUST 2024
	› Title : Data Augmentation based approach for Extremely Low Resource Machine Translation
	NLP Team, Google Research India, SEPTEMBER 2023
	› Title : Meta-Learning for Low-Resource Language Generation
	SCIS, University of Hyderabad, India, AUGUST 2023
	› Title : Large Language Models and Responsible Text Generation
	CSE Seminar Series, IIT Hyderabad, JUNE 2023
	› Title : Hype to Reality : Responsible Text Generation with Large Language Models
	SCIS, University of Hyderabad, India, APRIL 2022
	› Title : Introduction to Deep Learning through the Lenses of Natural Language Processing
	ASIC, NLP Reading and Discussion Group, MAY 2021
	› Title : An Unsupervised Cross-lingual Transfer Framework for Language Generation
	Faculty Development Program on Data Science, Hyderabad, India, DECEMBER 2020
	› Title : Introduction to RNN, Data Processing, and Text Classification
	ACM IIT Hyderabad Student Chapter, NOVEMBER 2020
	› Title : Learning to Distract : Generation of Incorrect Options from Reading Comprehension MCQs

⚙ RESEARCH SERVICES

Organizer	(1) Tutorial on "LMs in Education : Novel perspectives, challenges, and opportunities" at CO-LING 2025, Abu Dhabi.
	(2) One-day NLP meet on "Large-scale Human Evaluation for Indian Languages" at SCIS, UOH, 2016.
	(3) One-day workshop on "Sara Translator and Recent Trends in NLP" at SCIS, UOH, 2015.
Conference Session Chair	COLING 2025
Conference Reviewer	ACL [2021–24], EMNLP [2021–24], NAACL [2021-25], AAAI [2021–24], ARR Reviewer, ICLR 2021, SIGTYP 2021, MLR 2021
Student Volunteer	ACL [2020–22 (in-person)], EMNLP [2020–23], ICLR [2020–21], NeurIPS [2020–22], NAACL [2021–22], CIKM 2021, ACML 2022 (in-person)
Student Lead Volunteer	EMNLP 2022, ACML 2022

🏆 ACADEMIC HONORS & AWARDS

› Recipient of prestigious global **Google Academic Grant (GARA) 2024** as Co-PI on "*2σ Intelligent Tutoring System (ITS)*" ☑[Link].
› Received the prestigious **Annual Research Excellence Award at IIT Hyderabad** from the CSE department in 2024.
› Received a grant of 100K INR for attending the conference by IIT Hyderabad in the **Exceptional Research Scholar** category, 2022.
› Shared task **Best Paper Honorable Mention** for our CONSTRAINT workshop paper at AAAI 2021. ☑[Link]
› **Microsoft and ACL Travel Grant 2022 :** Received travel grants to present our Meta-XNLG work at ACL 2022 in Dublin, Ireland.
› Our **ZmBART** and **Meta-XNLG** papers were recognized as **Premier Papers** from India by IKDD-ACM India. ☑[Link]
› Received a **Student Travel Grant** to attend CODS-COMAD 2022 and 2023.
› Received **Fully-Funded Summer School Offers** from Oxford University ML Summer School in 2021, 2022, and 2023.
› Served as the **Head of Hybrid Infrastructure (Airmeet)** for the ACML Conference 2022.
› **ACM SIGIR and ACM India-IARCS Grant 2020 :** Received to present our CIKM 2020 work.
› **Suzuki Foundation Fellowship 2020 & 2021 :** Fully funded visit to Shizuoka University, Japan, for a short research stay.
› **BBC & IITG Hackathon :** Received a fully paid trip for a hackathon at Google's office in Gurugram, India, sponsored by BBC India and IITG.
› Participated in the **Google India Hackathon** at the Hyderabad office in 2018.
› **University of Tokyo and IITH Workshop 2018 :** Secured **second place** in the workshop (shared task) at IIT Hyderabad.
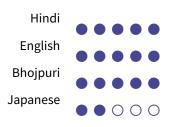› Qualified GATE in 2014, 2015, and 2016 with a **95+** percentile.

- Secured **second place** in high school in Ballia district, Uttar Pradesh, India.
- **Winning Captain** of National Cadet Corps (NCC) in 8th grade.

👥 TEACHING ASSISTANTSHIP AND MENTORSHIP

- Conducted **lectures and hands-on sessions** in the NLP/IR course at IITH during 2020–2024.
- **TA at MBZUAI :** NLP (Fall'24, 60+ students) and NLG (Spring'25, 15+ students).
- **TA at IITH :** NLP & IR (Spring'23, 45+ students), Information Retrieval (Fall'21, 25+ students), NLP (Spring'21, 110+ students), IR (Fall'20, 60+ students), NLP (Spring'20, 45+ students), IR (Fall'19, 30+ students), and Algorithms (Spring'19, 120+ students).
- **TA at UoH :** NLP (Fall'16, 20+ students) and Optimization Techniques (Spring'17, 55+ students).
- **Project Mentor at IITH :** NLP (Spring'24, mentored 2 groups), NLP & IR (Spring'23, mentored 4 groups), IR (Fall'21, mentored 7 groups), IR (Fall'20, mentored 8 groups), NLP (Summer'20, mentored 4 groups), and NLP (Winter'19, mentored 2 groups).
- Work on **Diverse Headline Generation** with Venkatesh E (MTech-TA at IITH) was published at Findings of ACL 2023.
- Work on **Hostility Detection in Code-Mixed Conversations** with Kamal Shrestha and Aditi Sanjay Bagora (MTech-TA students at IITH) was published at WebSci 2022.
- Work on **Coarse and Fine-Grained Hostility Detection** with Arkadipta De and Venkatesh E (MTech-TA students at IITH) was published at the CONSTRAINT workshop, AAAI 2021.

🈂 LANGUAGES

Hindi ● ● ● ● ●
English ● ● ● ● ●
Bhojpuri ● ● ● ● ●
Japanese ● ● ○ ○ ○

💡 INTERESTS

- **Love :** Traveling, Reading, Swimming, and Playing Strategic Games
- **Sports :** Cricket and Tennis