

INDIAN ACADEMY OF SCIENCES

Bengaluru

34th MID-YEAR MEETING

7 - 8 July, 2023

Venue: Auditorium, Biological Sciences Building, Indian Institute of Science, Bengaluru

Symposium: Developing ChatGPT for India: Challenges and opportunities in building large language models

Friday, 7 July 2023 - 14.30 to 16.30 h

14.30 - 14.35

Convener

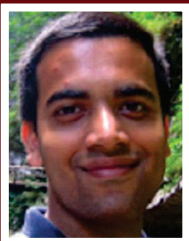


Chiranjib Bhattacharyya
IISc, Bengaluru

An overview:

Chiranjib Bhattacharyya is currently Professor and the chair of the Department of Computer Science and Automation, Indian Institute of Science. His research interests are in foundations of Machine Learning, Optimisation and their applications to Industrial problems. He has authored numerous papers in leading journals and conferences in Machine Learning. Some of his results have won the best paper awards.

14.35 - 14.55

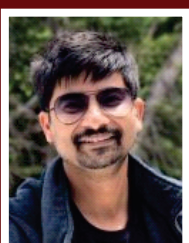


Partha Pratim Talukdar
CDS & CSA, IISc and Google Research India

Large language models for India

Partha Talukdar is a faculty member in the Department of Computational and Data Sciences (CDS) at the Indian Institute of Science (IISc), Bengaluru. He is also the founder of KENOME, an enterprise Knowledge graph company with the mission to help enterprises make sense of big dark data. Partha is interested in Machine Learning, Natural Language Processing, Data Integration, and Cognitive Science. His research has focused on learning and inference over graphs, automatic knowledge harvesting from web-scale datasets, temporal information processing, and neuro-semantic.

14.55 - 15.15



Monojit Choudhury
Microsoft Research India

Safety, infrastructure and language: The three scaling challenges of LLMs

Currently, he is a Principal Data and Applied Scientist at Turing India. They build large universal language models that form the backbone of various Microsoft products. He has a BTech and PhD in Computer Science and Engineering from IIT Kharagpur, and has been at Microsoft Research since 2007. He is interested in understanding the nature of the massively multilingual language models (such as Turing ULR, XLM-R, mBERT). As a part of Project LITMUS - Linguistically Aware Testing of Multilingual Systems, he worked on systematic evaluation and estimation of MMLM performance across languages, even in the absence of test datasets. This in turn enables them to understand the factors that affect MMLM performance and build optimal data collection strategies to ensure more equal or equitable performance.

15.15 - 15.35

Chairperson



Vasudeva Varma
IIIT Hyderabad and Brane Enterprises Pvt Ltd

Leveraging multilingual large language models to generate grounded encyclopaedic articles in Indian languages: Challenges and opportunities

Vasudeva Varma is the Head of Language Technologies Research Centre (LTRC), IIIT Hyderabad. He has served as the Dean (Research & Development), IIIT-H, and as the CEO of IIIT Hyderabad Foundation, one of the largest technology incubators in India that also managed IIIT-H's Intellectual Property and technology transfers. He has also co-founded Veooz Labs Pvt. Ltd, a social signal-based content discovery platform with social media search and analysis capabilities. He obtained his PhD from the Department of Computer and Information Sciences at the University of Hyderabad in 1996.

15.35 - 16.30

Panel Discussion / Q&A Session

Watch live on YouTube: <https://youtube.com/live/THe7oCrqbVk?feature=share>