MADAN MOHAN MALAVIYA UNIVERSITY OF TECHNOLOGY, GORAKHPUR

A GRADE
ACCREDITED BY
NAAC

ONLINE FACULTY DEVELOPMENT PROGRAMME

on

ADVANCED DEEP DIVE INTO THE
HIDDEN MATHEMATICAL SCIENCE OF MECHANISTIC INTERPRETABILITY

19th to 24th January 2026 | Online Mode

POWERED BY:
COMPUTER SCIENCE & ENGINEERING DEPARTMENT

MADAN MOHAN MALAVIYA UNIVERSITY OF TECHNOLOGY, GORAKHPUR

DEORIA ROAD, GORAKHPUR-273010 (UP) INDIA (U.P. Govt. Technical University) NAAC Grade 'A' University

University NIRF- Engineering: 60

PATRON:

Prof. Jay Prakash Saini

Hon'ble Vice Chancellor MMMUT Gorakhpur, India

Prof. Jay Prakash Saini, Hon'ble Vice Chancellor of Madan Mohan Malaviya University of Technology (MMMUT), Gorakhpur, serves as the Patron of this Online Faculty Development Programme on "Advanced Deep Dive into the Hidden Mathematical Science of Mechanistic Interpretability." A visionary academic leader committed to fostering research excellence and technological advancement, Prof. Saini continuously encourages innovation-driven learning and capacity building among faculty members. His insightful guidance and enduring support play a pivotal role in promoting high-quality academic initiatives and empowering the teaching community to engage with emerging frontiers in artificial intelligence and mathematical sciences.

CONVENER:

Dr. Rakesh Kumar

Head of Department Department of Computer Science & Engineering MMMUT Gorakhpur, India

Dr. Rakesh Kumar, Head of the Department of Computer Science and Engineering at Madan Mohan Malaviya University of Technology (MMMUT), Gorakhpur, is the Convener of this Online Faculty Development Programme on "Advanced Deep Dive into the Hidden Mathematical Science of Mechanistic Interpretability." A dedicated academician and proactive leader, Dr. Kumar strongly advocates research-driven learning and technological advancement. His focused guidance and commitment to academic excellence have been central to shaping this programme and enabling participants to explore emerging developments in artificial intelligence and mathematical sciences.

COORDINATORS

Dr. Satya Prakash Yadav

Associate Professor, Department of Computer Science & Engineering MMMUT Gorakhpur, India

Dr. Satya Prakash Yadav, Associate Professor in the Department of Computer Science and Engineering at Madan Mohan Malaviya University of Technology (MMMUT), Gorakhpur, is the Coordinator of this Online Faculty Development Programme on "Advanced Deep Dive into the Hidden Mathematical Science of Mechanistic Interpretability." A committed researcher and dedicated faculty member, Dr. Yadav actively contributes to fostering academic growth and advancing emerging technologies within the department. His meticulous planning and effective coordination have played a key role in shaping this programme and ensuring a valuable learning experience for all participants.

Dr. Shwet Ketu

Assistant Professor, Department of Computer Science & Engineering MMMUT Gorakhpur, India

Dr. Shwet Ketu is an Assistant Professor in the Department of Computer Science and Engineering at Madan Mohan Malaviya University of Technology (MMMUT), Gorakhpur. He holds a Ph.D. in Information Technology from Banaras Hindu University, Varanasi, and has previously served as an Assistant Professor at Galgotias University, UPES, and SIET. With a strong commitment to teaching and research, his work focuses on Data Analytics, supported by publications in reputable journals and conferences. Dr. Ketu is dedicated to creating a dynamic and inclusive learning environment that promotes critical thinking, creativity, and practical understanding, inspiring students to engage deeply with emerging technologies and real-world challenges.

ORGANIZERS

Dr. Satvik Vats

Assistant Professor, Department of Computer Science & Engineering MMMUT Gorakhpur, India

Dr. Satvik Vats (SMIEEE) is an Assistant Professor in the Department of Computer Science and Engineering at MMMUT, Gorakhpur, and holds a Ph.D. from BIT Mesra. With prior experience as an Assistant Professor at Graphic Era Hill University, Dehradun, and as an AI Researcher on the iDoc-X Project at Teleglobal Consulting Ltd., UK, he has contributed significantly to AI-driven healthcare solutions. Dr. Vats has numerous SCI/Scopus-indexed publications, authored books and book chapters, and holds 21 patents. A recipient of the Young Scientist Award (2025) and listed among the World's Top 2% Scientists (2025), he is also the Founder and Convener of the international conference series AutoCom and CyberCom. His work focuses on advancing research and innovation in Big Data Analytics, Machine Learning, and Deep Learning.

Dr. Shantanu Shahi

Assistant Professor, Department of Computer Science & Engineering MMMUT Gorakhpur, India

Dr. Shantanu Shahi is an Assistant Professor with an impressive academic background, holding a Post Doctorate from Malaysia and a Ph.D. in Image Recognition, along with M.Tech. and B.Tech. degrees in Computer Science and Engineering. His research interests span face recognition, deep learning, Deep Net architectures, FaceNet, and VGG models, with a strong emphasis on designing robust and high-accuracy recognition systems. Dr. Shahi's expertise extends to advanced algorithm development and IoT-based intelligent applications, where he focuses on creating practical, efficient, and scalable solutions for real-world challenges. Through his work, he continues to contribute to cutting-edge advancements in computer vision and intelligent systems.

About University

Madan Mohan Malaviya University of Technology (MMMUT), Gorakhpur, was established in 2013 by the Government of Uttar Pradesh as a non-affiliating state technical university. It was formed by reconstituting the Madan Mohan Malaviya Engineering College (MMMEC), Gorakhpur, which had been established in 1962. The University offers a wide range of academic programs, including Undergraduate (UG) courses in Civil Engineering, Chemical Engineering, Computer Science and Engineering, Mechanical Engineering, Electrical Engineering, Electronics and Communication Engineering, and Information Technology. In addition, the University offers B.Pharm, BBA, MBA, MCA, M.Tech, M.Sc., and Ph.D. programs in various specializations.

About Department

The Department of Computer Science and Engineering (CSE) at Madan Mohan Malaviya University of Technology, Gorakhpur, is one of the most dynamic and progressive departments of the University. Established in 1984, the department has evolved over the years into a center of academic excellence, offering comprehensive programs that blend theoretical foundations with practical and research-oriented learning.

The department offers Undergraduate (B.Tech in CSE), Postgraduate (M.Tech in CSE specialization), and Doctoral (Ph.D.) programs. The curriculum is regularly updated to keep pace with the rapidly evolving technology landscape and to meet global standards. Emphasis is placed on equipping students with the skills and knowledge required in emerging areas such as Artificial Intelligence, Machine Learning, Data Science, Cybersecurity, Cloud Computing, and Internet of Things (IoT).

TENTATIVE LIST OF SPEAKERS

- **Mr. Harish Singh**
  (Senior Manager, Accenture, Redmond, WA, USA)
- **Prof. Raghuraj Singh**
  (Professor, HBTU-Kanpur)
- **Prof. Dinesh Kumar Vishwakarma**
  (Professor, Delhi Technological University, Delhi)
- **Prof. Sonali Agarwal**
  (Professor, IIIT-Allahabad)
- **Mr. Nagendra Sharma**
  (Senior Director, Capgemini, North America)
- **Mr. Venkata Desai**
  (Senior Specialist, Microsoft Japan)
- **Mr. Deepak Bhalla**
  (Director, Estute Global Solutions Pty Ltd, Melbourne, Australia)
- **Mr. Dip Roy**
  (Director, WNS Global Service)
- **Prof. Dimitrios A. Karras**
  (National and Kapodistrian University of Athens (NKUA), Athens)
- **Dr. Allen Paul L. Esteban**
  (Assistant Professor IV, Nueva Ecija University of Science and Technology Cabanatuan City, Nueva Ecija)
- **Dr. Vibhash Yadav**
  (Associate Professor, Rajkiya Engineering College, Banda)
- **Dr. Sanjay Misra**
  (Senior Scientist, Institute for Energy Technology, Norway)

ABOUT FDP

This faculty development program aims to provide a comprehensive exploration of mechanistic interpretability, a key area of AI research dedicated to uncovering and understanding the underlying computational mechanisms and representations learned by neural networks. As AI systems become more advanced, gaining transparency into their decision-making processes is essential for ensuring safety, alignment, and trustworthiness in their deployment. Despite significant progress in the field, many challenges remain, including the need for more robust conceptual frameworks, improved methodologies for model analysis, and scalable strategies for aligning AI behavior with human values. This faculty development program presents state-of-the-art research, innovative methodologies, and real-world applications, offering insights into how mechanistic interpretability can help us understand, refine, and govern powerful AI systems. A key focus of this faculty development program is the integration of mechanistic interpretability with AI safety and alignment research. Traditional explainability techniques often rely on post-hoc interpretations, which can be limited in their ability to provide deep, mechanistic insights into neural network behavior. In contrast, mechanistic interpretability seeks to reverse-engineer AI models at a granular level, allowing researchers to develop more reliable, interpretable, and controllable AI systems. This faculty development program highlights cutting-edge research that demonstrates the potential of mechanistic approaches in improving transparency, interpretability, and AI governance.

TOPICS TO BE COVERED

- Basics of mechanistic interpretability and its role in AI
- Understanding deep neural networks through information theory
- Core mathematical tools for AI: linear algebra, calculus, and probability
- Concepts of sparse coding, Hebbian learning, and disentangling model features
- Techniques to visualize neural activations and interpret model attributions
- Bayesian and probabilistic approaches to interpreting neural models
- Mechanistic understanding of transformers and attention mechanisms
- Neurosymbolic methods and interpretable logic-based model layers
- Tools for identifying circuits and structures within large language models
- Global trends and future directions in interpretability research
- Interpretability challenges in federated and privacy-preserving AI.
- Ethics, explainability, and responsible AI governance

OBJECTIVES OF FDP

- Gain mathematical clarity on core mechanistic interpretability concepts.
- Learn practical tools and methods to analyze neural networks at different levels.
- Understand interpretability's role in AI safety and value alignment.
- Explore emerging trends in interpretability for large models, RL, and vision.
- Recognize ethical and governance issues linked to interpretability and its dual-use risks

ELIGIBILITY

- Faculty members of AICTE approved Education Institutions (Engineering and Polytechnic Colleges)
- Research Scholars , PG Scholars
- Industry Personnel

"NO CHARGE FOR REGISTRATION , COURSE AND CERTIFICATION"

REGISTRATION

Registration Procedure:
1. Sign up as a participant at: https://atalacademy.aicte-india.org/signup
2. FIll in your general details on the ATAL Portal
3. Upload your ID card and NOC/Registration form (PDF)
4. Navigate to FDPs, select type "ONLINE ATAL" , choose "January" under the month
5. Identify the FDP titled " Advanced Deep Dive into the Hidden Mathematical Science of Mechanistic Interpretability" starting from January 19th, 2026
6. Apply by clicking the '+' icon. A confirmation message will be displayed, and the FDP will appear under the "Applied FDPs" section.

TO REGISTER SCAN HERE ➤

CONTACT FOR ANY QUERY

- **Dr. Satya Prakash Yadav**
  (Associate Professor, Department of Computer Science & Engineering MMMUT Gorakhpur, India)
  Email: spycs@mmmut.ac.in **Mob:** +91-7678647454

- **Dr. Shwet Ketu**
  (Assistant Professor, Department of Computer Science & Engineering MMMUT Gorakhpur, India)
  Email: skcse@mmmut.ac.in Mob: +91-8175848784

- **Dr. Satvik Vats**
  (Assistant Professor, Department of Computer Science & Engineering MMMUT Gorakhpur, India)
  Email: svcse@mmmut.ac.in **Mob:**+91-7428262793

- **Dr. Shantanu Shahi**
  (Assistant Professor, Department of Computer Science & Engineering MMMUT Gorakhpur, India)
  Email: dsscse@mmmut.ac.in Mob: +91-9026111993.