National Institute of Technology Rourkela

Progress Seminar

Seminar Title	: Effective Human-Robot interaction with Multimodal Data Analysis
Speaker	: Sougatamoy Biswas (Rollno : 521cs6015)
Supervisor	: Anup Nandy
Venue	: CS323 (CSE Department)
Date and Time	: 04 Sep 2025 (05:15 PM)
Abstract	: Human robot interaction (HRI) is a rapidly growing field of research, with the goal of developing robots that can interact with humans in a natural and effective way. Multimodal HRI is a subfield of HRI that focuses on the use of multiple communication channels, such as speech, gesture, and touch, to improve the efficiency and effectiveness of human-robot interaction. Deep learning techniques have significantly enriched this field by enabling robots to perceive and respond to diverse communication cues, including speech, gestures, and facial expressions. Fusing information from multiple modalities allows robots to gain a deeper understanding of user intentions, context, and emotions, leading to more intuitive and effective interactions. This research aims to explore multiple modalities and their fusion strategies, encompassing real-time gesture recognition, emotion recognition, and activity detection. This study provides a novel approach for different modalities that can be used for multimodal HRI. In the current analysis, the objective is to develop robots that can understand and respond to human communication in a natural and effective way. This can be achieved by using multiple communication channels, such as speech, gesture, and touch. Deep learning techniques are employed to extract features from multiple communication channels, which are then fused to represent human intent. This representation enables the robot to interpret human communication and respond effectively. A comparative evaluation is conducted to assess the effectiveness of the proposed multimodal fusion strategy, which demonstrates improved accuracy over existing methods. This study also outlines recent developments, identifies current limitations, and suggests future directions in multimodal human-robot interaction using deep learning.