



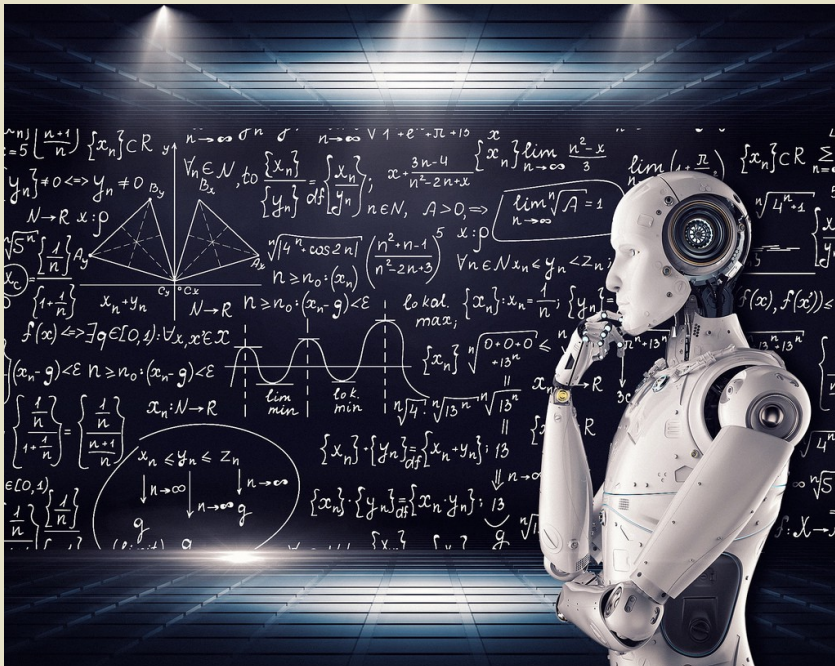
May 2022 | Volume 1 | Edition 9

# AIML INSHORTS

Department of Artificial Intelligence and Machine Learning

## CONTENTS

2. Department Vision and Mission
3. AI Realm
4. Reach Us



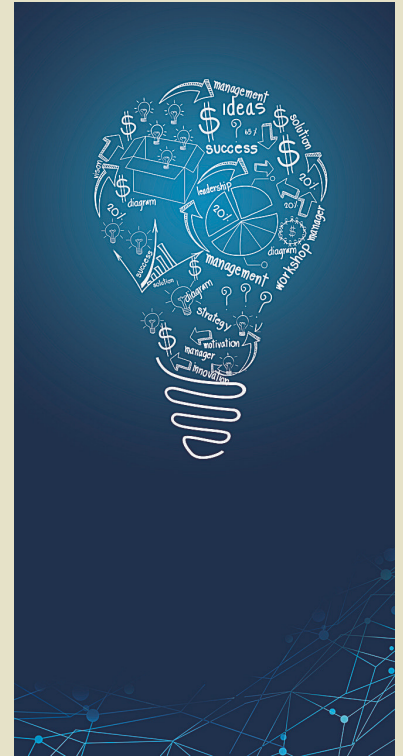
**“Machine learning and deep learning will create a new set of hot jobs in the next 5 years.” -Dave Waters**



# DEPARTMENT VISION AND MISSION

## VISION

To be a department focused on quality education and research in Artificial Intelligence & Machine Learning that prepares early professionals contributing to serve the society.

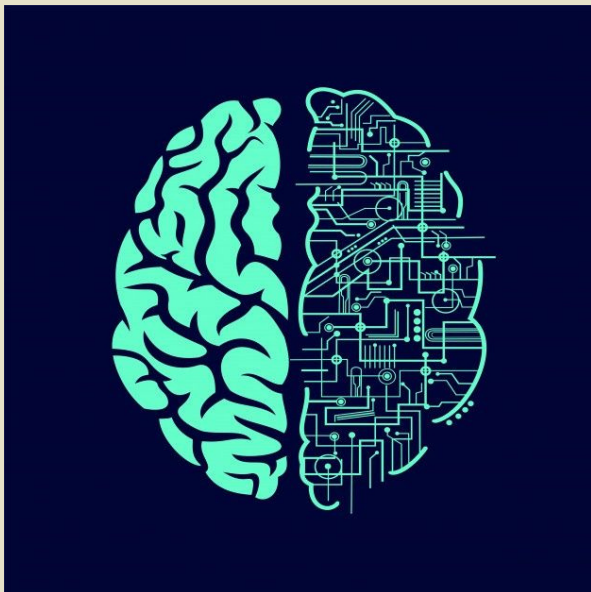


## MISSION

To provide an academic environment for the development of professionals in the field of Artificial Intelligence and Machine Learning.

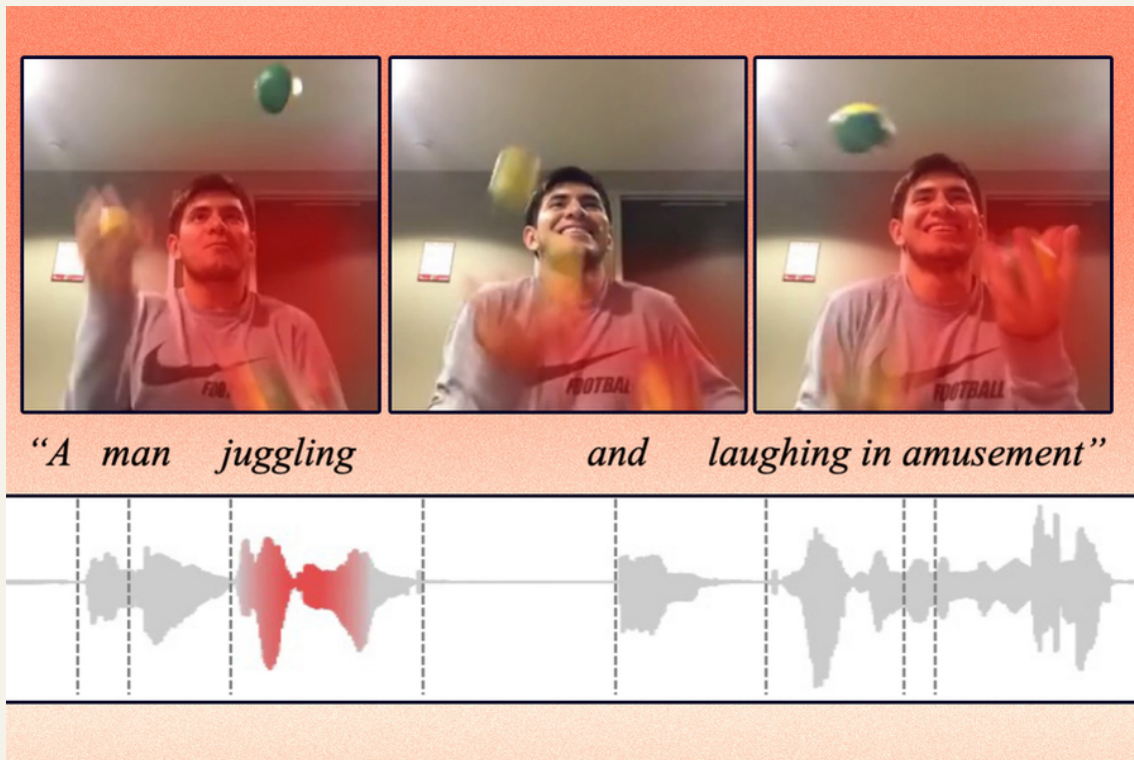
To cultivate research culture resulting in knowledge and development of the society.

To enhance academic collaborations for better exposure.





## A MACHINE-LEARNING MODEL CAN IDENTIFY THE ACTION IN A VIDEO CLIP AND LABEL IT, WITHOUT THE HELP OF HUMANS.



Liu and his collaborators developed an artificial intelligence technique that learns to represent data to capture concepts shared between visual and audio modalities. For instance, their method can learn that the action of a baby crying in a video is related to the spoken word "crying" in an audio clip.

Using this knowledge, their machine-learning model can identify where a certain action is taking place in a video and label it.

It performs better than other machine-learning methods at cross-modal retrieval tasks, which involve finding a piece of data, like a video, that matches a user's query given in another form, like spoken language. Their model also makes it easier for users to see why the machine thinks the video it retrieved matches their query.

This technique could someday be utilized to help robots learn about concepts in the world through perception, more like the way humans do.

The representation learning model takes raw data, such as videos and their corresponding text captions, and encodes them by extracting features, or observations about objects and actions in the video. Then it maps those data points in a grid, known as an embedding space. The model clusters similar data together as single points in the grid. Each of these data points, or vectors, is represented by an individual word.



# REACH US



[ucoe.edu.in](http://ucoe.edu.in)



Near Bhajansons & Punyadham, Kaman Bhiwandi Road, Vasai, Dist. Palghar - 401 205



+ 91 8407979167 / 8007090722 / 8007478659



[info.ucoe@universalcollege.edu.in](mailto:info.ucoe@universalcollege.edu.in)



UCoEKaman



ucoe\_mumbai

