NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

Understanding Machine Learning and Computer Vision

Introduction

This course will introduce important fundamental concepts and theories in machine learning and computer vision from intuitive ideas and common senses. Artificial Intelligence (AI) is one of the most important technologies in recent years. Industrial demand for AI engineers, researchers and professionals is exploding. While computer vision has showed the most widely successful and impressive applications of AI, machine learning is the most important and effective solution to current AI problems. However, some abstract concepts and complex theories hinder many people from getting acquainted with the fields of machine learning and computer vision. This course will start from common senses and intuitive ideas that we know, and identify the limitations of these ordinary ideas and find the natural ways to overcome them, while introducing important fundamental concepts and theories in machine learning and computer vision in an easy-to-understand way. In particular, the course will cover from fundamental theories of machine learning to the state-of-the-art solutions, such as Convolutional Neural Networks (CNNs) and the Transformers. At the end of the course, students will have a good understanding of the important fundamental concepts and theories in machine learning and computer vision, and appreciate that AI is not a black box, but it is explainable.

Course Objectives

This course aims at providing students with a good understanding of the abstract concepts and fundamental theories in machine learning and computer vision. Students will discover that there is neither miracle nor mystery in AI, machine learning and computer vision. All abstract concepts and complex theories in AI, machine learning and computer vision have been developed from intuitive ideas and common senses that everyone can understand. This course aims at clearing the barriers that may prevent you from understanding the topics of AI, machine learning and computer vision. It will also equip you with the capability to understand advanced concepts in the relevant fields.

Course outline

Session 1: Introduction to Visual Object Recognition, Intuitions of Machine Learning

Intuitive understanding of visual object recognition in computer vision: from rule-based approach to real-example comparison, and from simple comparison to statistical classification. The critical and important role of normalization in visual object recognition and machine learning. Understand why machine learning is necessary to solve recognition problems.

Session 2: From Decision and Classification with Probability Theory to Machine Learning

Understand why probability theory is necessary and the foundation of recognition. Intuitive understand and visualization of probability theory and methods. The probabilistic representation of decision, prediction, and classification. Understand what is machine learning? Kernel function and nearest neighbor approaches for machine learning.

Session 3: Fundamentals of Machine Learning and Neural Networks

Maximum likelihood (ML) learning and minimum mean square error learning. From probability as objective function to expectation (mean, average or summation) of errors as loss function. Concepts, structure and training of the traditional fully connected neural networks (MLPs). The capabilities and limitations of traditional fully connected neural networks (MLPs).

Session 4: Deep Learning: from MLP to CNN and from CNN to the Transformer

The concepts of deep learning and the principles of convolutional neural networks (CNNs). How CNN solves problems of the traditional fully connected neural networks. The capabilities and limitations of CNNs in computer vision. Form limitations of CNN to the Transformer. Concepts and principles of Transformer. How Transformer overcomes the limitations of CNN. CNN solves problems of MLP by deviating from global to local. Why does transformer attend to global yet outperform CNN in general? The relation between Transformer and CNN and the difference between Transformer to MLP.

Session 5: Group Presentation by Students, Discussion and Question Answer

Presentation of students' view and understanding to a concept per group of students. Each group of students chooses a concept they gain the most understanding in the course. Further deepen the full understanding to the concept by discussion and question answer. Evaluation and the Best Presenting Team selection.

Duration

3 hours x 5 sessions= 15 hours

Teaching Methodology

In class lecturing, discussion and question answer, Group Presentation by students.

About the Instructor

Dr Xudong Jiang (Fellow, IEEE) received the B.E. and M.E. degrees from the University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 1983 and 1986, respectively, and the Ph.D. degree from Helmut Schmidt University, Hamburg, Germany, in 1997. From 1986 to 1993, he was a Lecturer with UESTC. From 1998 to 2004, he was with the Institute for Infocomm Research, A*STAR, Singapore, as a Lead Scientist, and the Head of the Biometrics Laboratory, Singapore. In 2004, he joined Nanyang Technological University (NTU), Singapore, as a Faculty Member, where he served as the Director of the Centre for Information Security from 2005 to 2011. He is currently an Associate Professor with the School of EEE, NTU, where he is also the Director of the Centre for Information Sciences and Systems and serves as Deputy Director of MSc Signal Processing and Machine Learning Program. He has authored more than 200 articles with more than 60 articles in IEEE journals, including 10 articles in IEEE Transactions on Pattern Analysis and Machine Intelligence and 18 articles in IEEE Transactions on Image Processing. Recent years, he published over 20 papers in top computer vision conferences such as CVPR, ICCV, ECCV and AAAI. His current research interests include pattern recognition, computer vision, machine learning and image processing. Dr. Jiang has served as an Associate Editor for IEEE Signal Processing Letters and IEEE Transactions on Image Processing. He is currently serving as a Senior Area Editor for IEEE Transactions on Image Processing and the Editor-in-Chief for IET Biometrics.