Speaker: Dr. V. Ramasubramanian, Professor, IIIT Bangalore

Title: Cybenko's universal approximation theorem: The corner stone of deep-learning.

Abstract:
Presently, 'deep learning' theory and practice can be considered as one of the key elements transforming the 'IT' industry's landscape (specifically referring to the big ones – Google, Apple, Microsoft, IBM, Amazon etc.) which has product-spaces that attempt to automate human cognitive abilities, hinged on solving a diverse set of "AI problems" such as automatic speech recognition (ASR), machine vision, sequence classification, natural language processing (NLP), machine translation (MT), recommendation systems, advanced driver assistance systems (ADAS) etc. 'Deep learning', even while being a catchy buzz-word on one hand, is indeed to be acknowledged as a phenomenal breakthrough in "machine-learning", forming the core of the emerging trends highlighted above.
However, while having caught the fancy of the theorists and practitioners alike, what is far less known is that the several variants of 'deep learning' neural architectures, notably the 'deep neural network' (DNN), owe their ability (to form viable solutions to the various problem domains listed above), largely to a fundamental result called the "Universal Approximation Theorem' by Cybenko in the late 80s [1]. Cybenko's theorem establishes the approximation power of superpositions of a sigmoidal function, specifically showing that "arbitrary decision regions can be arbitrarily well approximated by continuous feedforward neural networks with only a single internal, hidden layer and any continuous sigmoidal nonlinearity", which in turn enables it to generate non-linear decision surfaces needed to handle linearly non-separable multi-class classification problems – a key result which forms the basis of and extends to the general case when the DNN is 'deep', i.e. having 'many' hidden layers as required for real-world problems. No wonder Eugene Wigner (Nobel Laureate Physicist) remarked about "The Unreasonable Effectiveness of Mathematics in the Natural Sciences" [2] which clearly extends to AI problems too.
This talk will walk through Cybenko's result, and its role in DNNs that form a class of 'deep learning' mechanisms.
[1] George Cybenko, "Approximation by superpositions of a sigmoidal function", Mathematics of Control, Signals, and Systems (MCSS), Vol. 2, Issue 4, (303-314), 1989, Springer-Verlag, NY.
[2] Eugene Wigner, "The Unreasonable Effectiveness of Mathematics in the Natural Sciences," in Communications in Pure and Applied Mathematics, vol. 13, No. I (February 1960). New York: John Wiley & Sons, Inc. Copyright © 1960 by John Wiley & Sons, Inc.