From MoE to MoD as well as MoX, Bi-directional intelligence, and Search Reasoning：Some Review and View

Lei Xu
Shanghai Jiao Tong University, China

Deepseek comes to be a very hot LLM topic in the recent two months, using the mixture of experts(MoE) architecture to coordinate many relatively small deep neural networks with the help of MHA,MTP,RL reasoning. By nature, it is a mixture of deep learning (MoD) and an exemplar of MoX, where X indicates any model for a specific task. This talk is about some review and view on AI advances along the line of MoX, starting from Hough transform to learning mixture of lines and Gaussian models, and further to MoE and mixture of generative or diffusion models, as well as bi-directional intelligence for AlphaGo-like problem solving and search reasoning, together with some views on their links to ChatGPT-4 and Deepseek-R1.