Analysis Seminar

Mathematical Foundations of DeepSeek: A Reinforcement Learning Perspective.

By

Berkay Anahtarcı
(Özyeğin University)

Abstract: This talk explores the mathematical foundations of DeepSeek R1, an RL-only model designed for complex reasoning. Unlike traditional supervised fine-tuning, DeepSeek R1 leverages Group Relative Policy Optimization (GRPO), a novel method that stabilizes Proximal Policy Optimization (PPO) without a critic. GRPO enhances chain-of-thought reasoning by structuring problem-solving into sequential steps. I will analyze its theoretical properties and implications for reasoning-driven reinforcement learning.

Date: Monday, February 24, 2025
Time: 17:00 – 18:00, GMT+3
Place: ZOOM
This is an online seminar. To request the zoom link, please send a message to goncha@fen.bilkent.edu.tr