

ÖDEV KAĞIDI 3

Model eğitim verisini aşırı temsil ettiğinde genelleme kabiliyeti düşer. Veri setini büyütmek ve regularization uygulamak etkili çözümlerdir.

Exploration, bilinmeyen eylemleri keşfetmek için gereklidir; exploitation ise mevcut bilgiyi optimize eder. İkisinin dengesi performansı belirler.

Transformer mimarisi sıralı bağımlılığı ortadan kaldırır ve GPU üzerinde verimli çalışır.