

ÖDEV KAĞIDI 18

Aşırı öğrenme, modelin ezber yapmasına yol açar. Bu durum çapraz doğrulama, dropout ve model karmaşıklığını azaltma yöntemleriyle kontrol altına alınabilir.

Yeterli keşif yapılmazsa ajan yanlış politikaya saplanabilir; aşırı keşif ise kararsızlığa yol açar. Bu nedenle denge kritik önemdedir.

Attention tabanlı yapı sayesinde Transformerlar uzun dizilerde bilgi kaybı yaşamaz ve paralel hesaplama avantajı sunar.