

Makale Özeti 1

Ilias Mantziaras

Mart 2025

1 Kullanılan Makale

Revisiting Transformer-based Models for Long Document Classification 2022
arXiv

2 Özet

Bilgisayarlar insanların yazdıklarını, hikayeleri veya mesajları anlayabilir. Ancak bazen bu yazılar çok uzundur, bir kitap gibi. Transformers adı verilen özel bilgisayar programları bu uzun yazıları anlamamıza yardımcı olabilir, ancak yazı çok uzunsa kafanız karışabilir. Bunu düzeltmek için insanlar daha uzun yazıları işleyebilen yeni Transformers sürümleri yaptı. Bu yeni sürümlere Longformer ve Hierarchical Transformers adı verilir. Bilgisayarın yazıyı küçük parçalara ayırarak veya küçük parçalara bölerek anlamasına yardımcı olurlar. Bu, bilgisayarın yazının ne hakkında olduğunu anlamasını kolaylaştırır.

2.1 Modeller

Uzun belgeler iki model kullanılarak işlenebilir: Longformer ve hiyerarşik Transformer'lar. Longformer, küresel dikkat için az sayıda belirteç kullanır ve bu da eğitim sürecini daha kararlı hale getirir. Hiyerarşik Transformer'lar bir belgeyi daha küçük parçalara böler ve bunları ayrı ayrı işler. Her veri kümesi için en uygun parça uzunluğu farklıdır. Belgeleri örtüşen parçalara bölmek, modelin daha iyi çalışmasına yardımcı olabilir. Etiket Bilinçli Dikkat Ağı (LWAN) eklemek de modelin performansını iyileştirebilir. LWAN, modelin her etiket için farklı gösterimler öğrenmesini sağlar. Longformer ve hiyerarşik Transformer'lar, yalnızca 512 belirtece kadar işleyen modellerden daha uzun metinleri işleyebilir ve daha iyi sonuçlar elde edebilir. Hiyerarşik modeller, Longformer'dan daha az GPU belleği kullanır. Hiperparametreleri ayarlamak istiyorsanız, başlangıç noktası olarak Longformer'ı kullanmanız ve hiyerarşik Transformer'ları göz önünde bulundurmanız önerilir.

2.2 Transformer'lar

Transformer'lar, belgeler gibi uzun yazı parçalarını anlamamıza yardımcı olabilen özel bilgisayar programlarıdır. Bazı insanlar Transformer'ların uzun belgeleri işleyemeyeceğini düşünmüşlerdir, ancak bu çalışma aslında işleyebildiklerini göstermektedir. Çalışmada farklı türde Transformer'lar kullanılmış ve tıbbi notlar ve hukuki davalar gibi farklı türdeki belgeler üzerinde test edilmiştir. Sonuçlar Transformer'ların belgeleri anlayabildiğini ve hatta bazı durumlarda diğer bilgisayar programlarından daha iyi performans gösterdiğini göstermiştir. Bu iyi bir haber çünkü Transformer'ların uzun belgeleri okuma ve anlama gibi birçok farklı görevde bize yardımcı olmak için kullanılabileceği anlamına geliyor.