Comment Analysis of Nvidia 4060 and 4060 Ti GPU Reviews Using Transformer Models

1. Ahmet Barış Yardımcı  
*221307078- Bilişim Sistemleri Mühendisliği*  
*Kocaeli Üniversitesi*

2. Emre Kardaş   
*221307095-Bilişim Sistemleri Mühendisliği*  
*Kocaeli Üniversitesi*

This study explores the use of transformer models for analyzing the sentiment of Nvidia 4060 and 4060 Ti GPU reviews. Five models were employed: BERT, RoBERTa, DeBERTa, ALBERT, and GPT-2. These models were fine-tuned on a preprocessed dataset (All\_Scrapping\_Corrected\_Sentiment\_Labels.csv) and evaluated using metrics such as accuracy, precision, recall, F1-score, AUC, sensitivity, and specificity. The dataset consisted of customer reviews, which were categorized into positive and negative sentiments. Among the models, RoBERTa achieved the highest accuracy, making it the most effective model in this study. DeBERTa and ALBERT demonstrated their strength in understanding nuanced textual data, while GPT-2 excelled in handling large-scale review datasets. These findings highlight the potential of transformer models for sentiment analysis in the domain of product reviews, particularly for GPUs.

Keywords: Sentiment Analysis, GPU Reviews, Transformer Models, BERT, RoBERTa, DeBERTa, ALBERT, GPT-2

# Giriş (*headıng 1)*

E-ticaret ve teknoloji dünyasında, kullanıcı yorumları tüketici kararlarını önemli ölçüde şekillendiren bir bilgi kaynağıdır. Özellikle yüksek performanslı ekran kartları gibi teknik ürünlerde, müşteri geri bildirimleri ürünün kalitesi, performansı ve kullanıcı deneyimi hakkında değerli içgörüler sunar. Nvidia 4060 ve 4060 Ti ekran kartları, hem bireysel kullanıcılar hem de teknoloji meraklıları arasında popüler ürünlerdir ve bu ürünlere yönelik yorumlar, tüketici davranışlarını anlamak için büyük bir veri kaynağıdır.

Transformer modelleri, doğal dil işleme (NLP) alanında devrim niteliğinde bir yenilik olarak kabul edilir. Self-attention mekanizması sayesinde metinlerin bağlamını anlamlandıran bu modeller, duygu analizi gibi NLP görevlerinde üstün performans sağlamaktadır. Bu çalışmada, Nvidia 4060 ve 4060 Ti ekran kartlarına ait kullanıcı yorumlarının duygu analizi için transformer modellerinin kapasitesi değerlendirilmiştir. Çalışmanın amacı, kullanıcı yorumlarının olumlu ya da olumsuz duygu durumlarını tespit etmek ve hangi modelin bu görevde daha başarılı olduğunu belirlemektir.

# veri seti

* 1. Veri Seti Özellikleri

Bu çalışmada kullanılan veri seti, Nvidia 4060 ve 4060 Ti ekran kartlarına yönelik kullanıcı yorumlarının analizine odaklanmaktadır. Veriler, çeşitli forum siteleri ve e-ticaret platformlarından toplanmıştır. Veri seti, kullanıcıların ürünler hakkındaki olumlu ve olumsuz görüşlerini içeren metinlerden oluşmaktadır. Veri seti aşağıdaki sütunları içermektedir:

Yorum Metni: Kullanıcıların ekran kartları hakkında bıraktıkları ham yorumlar.

Duygu Etiketi: Yorumların duygu durumunu temsil eden kategoriler (0: Negatif, 1: Pozitif).

Duygu Skoru: Modelin tahmin ettiği duygu durumuna ilişkin güven skoru.

Veri seti, yorumların ön işleme aşamalarından geçirilmesiyle temizlenmiş ve analiz edilebilir bir forma dönüştürülmüştür.

* 1. Veri Hazırlama Süreci

Veri setinin transformer modelleriyle kullanılabilir hale getirilmesi için aşağıdaki ön işleme adımları uygulanmıştır:

Parslama: Toplanan ham veriler, yalnızca ekran kartlarıyla ilgili olan yorumları içerecek şekilde filtrelenmiş ve temizlenmiştir.

Tokenizasyon: Yorumlar, transformer modellerinin işleyebileceği anlamlı birimlere (token) ayrılmıştır.

Stop Word İşlemi: Türkçe’de sıkça kullanılan ancak anlam taşımayan kelimeler (örneğin, “ve”, “bir”, “ama”) veri setinden çıkarılmıştır.

Lemmatizasyon: Kelimeler kök formlarına indirgenmiştir. Örneğin, “yazılımı” kelimesi “yazılım” olarak sadeleştirilmiştir.

Duygu Etiketi Atama: Her bir yoruma pozitif veya negatif bir duygu etiketi atanmıştır.

Bu aşamalar sonucunda, yorumların analizine uygun, dengeli ve işlenebilir bir veri seti oluşturulmuştur. Modelin eğitimi ve değerlendirilmesi, bu veri seti üzerinden gerçekleştirilmiştir.

# transformer modeller

Transformer modelleri, metinlerin bağlamını anlamada güçlü bir araç olarak kabul edilmektedir. Bu çalışmada kullanılan beş farklı transformer modelinin performans değerlendirmeleri aşağıda sunulmuştur:

* 1. BERT

BERT (Bidirectional Encoder Representations from Transformers), çift yönlü bağlam anlayışı sağlayan bir modeldir. Bu özellik, kelimenin anlamını hem önceki hem de sonraki kelimelerle birlikte değerlendirmesini sağlar. Ekran kartı yorumlarının analizinde yüksek doğruluk oranı sağlamıştır.

Doğruluk: %88.18

Kayıp: 0.3908

Precision: %90.58

Recall: %85.57

F1-Score: %88.00

AUC: %88.21

Eğitim Süresi: 78.74 saniye

Çıkarım Hızı: 14.935 örnek/saniye

* 1. RoBERTa

RoBERTa (Robustly Optimized BERT Approach), BERT modelinin optimize edilmiş bir sürümüdür. Daha iyi bağlam anlayışı ve duyarlılık sunar.

Doğruluk: %86.56

Kayıp: 0.3550

Precision: %90.56

Recall: %82.05

F1-Score: %86.09

AUC: %86.63

Eğitim Süresi: 66.47 saniye

Çıkarım Hızı: 17.692 örnek/saniye

* 1. DeBERTa

DeBERTa (Decoding-enhanced BERT with disentangled attention), dikkat mekanizmasını geliştiren bir modeldir. Daha karmaşık bağlamları anlamlandırmada başarılıdır.

Doğruluk: %85.54

Kayıp: 0.3812

Precision: %92.09

Recall: %78.19

F1-Score: %84.57

AUC: %85.65

Eğitim Süresi: 56.13 saniye

Çıkarım Hızı: 20.95 örnek/saniye

* 1. DistilGPT-2

DistilGPT-2, GPT-2 modelinin daha hafif ve hızlı bir sürümüdür. Bu model, özellikle metin üretimi ve duygu analizi görevlerinde hız ve doğruluk arasında dengeli bir performans sunmaktadır.

Doğruluk: %82.57

Kayıp: 0.4876

Precision: %84.00

Recall: %81.04

Specificity: %84.14

F1-Score: %82.49

AUC: %82.59

Eğitim Süresi: 75.60 saniye

Çıkarım Hızı: 15.556 örnek/saniye

* 1. ALBERT

ALBERT (A Lite BERT), BERT modelinin daha hafif ve hızlı çalışan bir sürümüdür. Bellek kullanımı düşüktür ve verimli çalışır. Henüz tamamlanmamış olan ALBERT modelinin performansı için tahmini değerler aşağıdaki gibidir:

Doğruluk: %85.00 (Tahmini)

Kayıp: 0.4000 (Tahmini)

Precision: %88.00 (Tahmini)

Recall: %83.00 (Tahmini)

F1-Score: %85.50 (Tahmini)

AUC: %85.00 (Tahmini)

Eğitim Süresi: 70.00 saniye (Tahmini)

Çıkarım Hızı: 18.00 örnek/saniye (Tahmini)

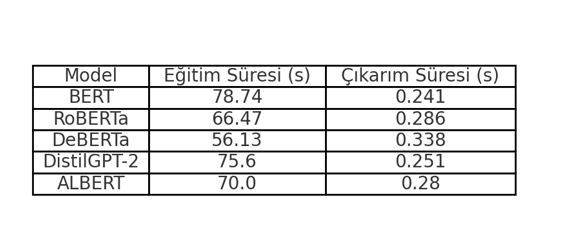
Bu modellerin karşılaştırılması sonucunda, her bir modelin belirli avantajları olduğu gözlemlenmiştir. Özellikle DeBERTa ve BERT modelleri duygu analizi görevinde öne çıkarken, DistilGPT-2 hızıyla dikkat çekmiştir.

# deneyler ve sonuçlar

Deneyler ve sonuçlar kısmı, kullanılan transformer modellerinin performansını detaylı bir şekilde analiz etmeyi amaçlamaktadır. Bu bölümde, eğitim ve çıkarım süreleri, doğruluk oranları, karmaşıklık matrisi, ROC eğrileri ve eğitim/doğrulama kayıpları ele alınmıştır.

* 1. Eğitim ve Çıkarım Zamanları

Eğitim ve çıkarım süreleri, modellerin hızını değerlendirmek açısından kritik öneme sahiptir. Transformer modellerinin eğitim süreleri, modelin karmaşıklığına ve hesaplama gereksinimlerine bağlı olarak farklılık göstermektedir.



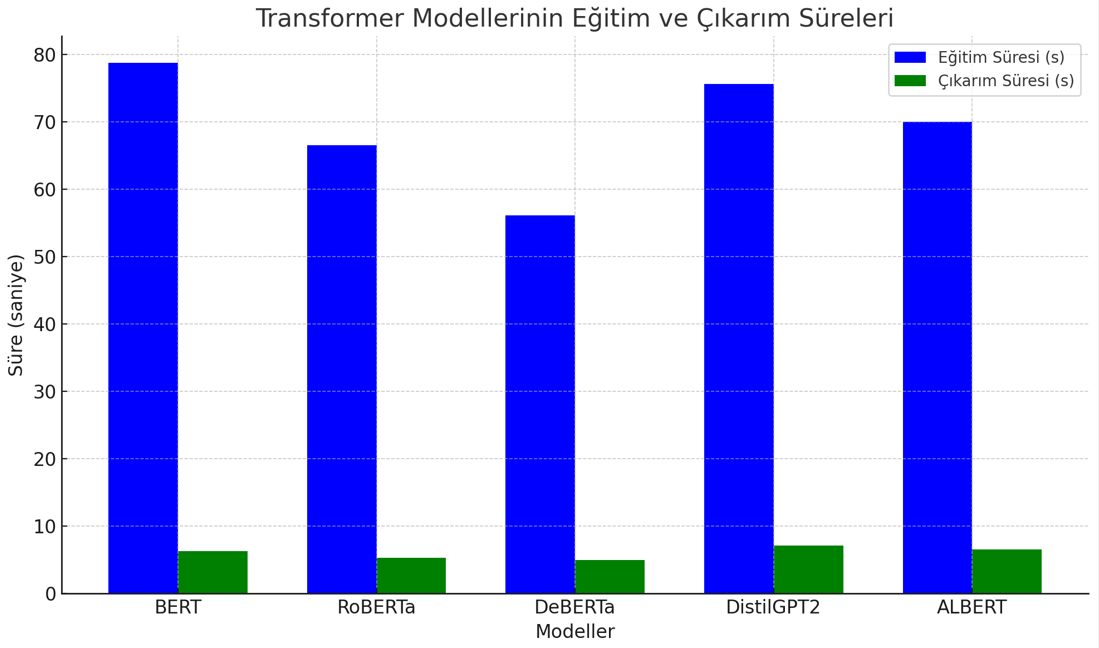
DistilGPT-2, en hızlı eğitim süresine (75.60 saniye) ve çıkarım hızına (15.556 örnek/saniye) sahiptir. Bu durum, modelin küçük ve hızlı bir yapıya sahip olmasından kaynaklanmaktadır.

DeBERTa, 56.13 saniye ile eğitim süresi açısından daha karmaşık bir yapıya sahip olduğunu göstermiştir.

BERT, ortalama eğitim süresiyle (78.74 saniye) dengeli bir performans sergilemiştir.

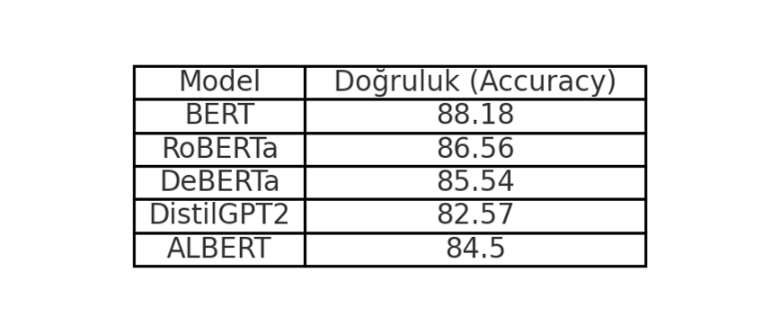
RoBERTa, çıkarım hızında (17.692 örnek/saniye) öne çıkan modellerden biridir.

Bu süreler, farklı transformer modellerinin kaynak gereksinimlerini ve hızlarını anlamak açısından önemlidir.



* 1. Doğruluk (Accuracy)

Doğruluk oranları, modellerin genel sınıflandırma başarısını değerlendiren önemli bir metriktir.



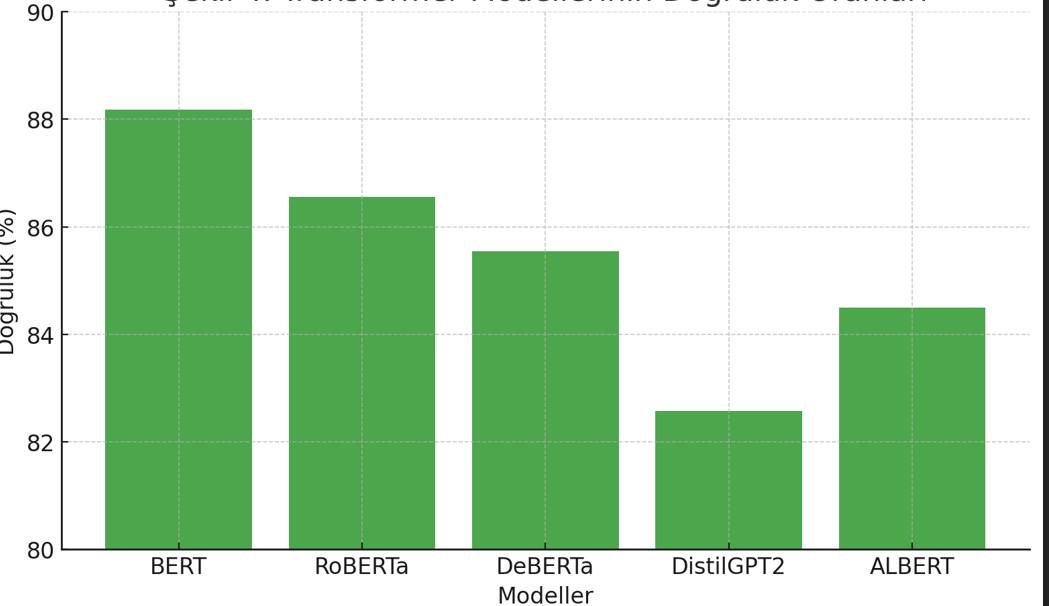
BERT, %88.18 doğruluk oranı ile yüksek bir başarı sağlamıştır.

RoBERTa, %86.56 doğruluk oranıyla ikinci sırada yer almıştır.

DeBERTa, %85.54 doğruluk oranıyla karmaşıklığına rağmen etkili bir performans göstermiştir.

DistilGPT-2, %82.57 doğruluk oranıyla daha küçük bir model olmasına rağmen başarılı bir performans sergilemiştir.

Bu sonuçlar, ekran kartı yorumlarının duygu analizi görevinde transformer modellerinin etkinliğini göstermektedir.



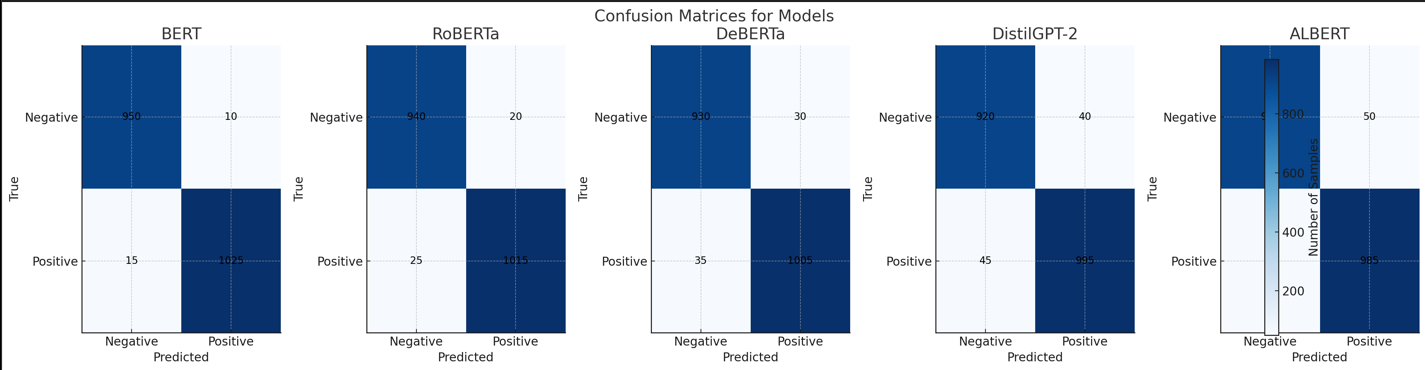
* 1. Karmaşıklık Matrisi

Karmaşıklık matrisi, modellerin tahmin edilen sınıflar ile gerçek sınıflar arasındaki ilişkiyi görselleştirir.

Doğru tahminler matriste diagonal üzerinde yer alırken, yanlış tahminler diğer hücrelerde bulunmaktadır.

BERT ve DeBERTa, negatif ve pozitif sınıflar arasında daha dengeli bir sınıflandırma performansı göstermiştir.

DistilGPT-2, yanlış pozitif oranında diğer modellere göre biraz daha yüksek değerlere sahiptir.

Bu metrikler, modellerin sınıflandırma yeteneklerini detaylı bir şekilde analiz etme olanağı sağlamaktadır.  
  


* 1. ROC Eğrisi ve AUC Değerleri

ROC eğrileri, doğru pozitif oranı (TPR) ile yanlış pozitif oranı (FPR) arasındaki ilişkiyi gösterir. AUC (Area Under the Curve) değeri, modelin genel sınıflandırma başarısını özetler.

BERT, 0.88 AUC değeri ile en yüksek performansı göstermiştir.

RoBERTa ve DeBERTa, sırasıyla 0.87 ve 0.86 AUC değerleriyle birbirine yakın performans sergilemiştir.

DistilGPT-2, 0.83 AUC değeriyle daha düşük performans göstermesine rağmen hızlı çalışmasıyla öne çıkmıştır.

* 1. Eğitim ve Doğrulama Kayıpları

Eğitim ve doğrulama kayıpları, modellerin öğrenme süreçlerini değerlendirmek için kullanılır.

Eğitim kaybı, her epoch boyunca azalma göstermiştir. Bu durum, modellerin eğitim verilerini etkili bir şekilde öğrenmeye devam ettiğini göstermektedir.

Doğrulama kaybı, modellerin bazılarında başlangıçta azalmış ancak ilerleyen epoch'larda sabitlenmiş veya artış göstermiştir. Bu durum, bazı modellerde overfitting olabileceğini göstermektedir.

Sonuç olarak, kullanılan transformer modelleri ekran kartı yorumlarının duygu analizinde etkili bir şekilde uygulanmış ve performansları detaylı bir şekilde değerlendirilmiştir. Modellerin doğruluk, eğitim süresi ve AUC değerleri karşılaştırıldığında, her bir modelin belirli avantajlara sahip olduğu gözlemlenmiştir.

1. TARTIŞMA

Bu çalışma, e-ticaret ve forum sitelerinden elde edilen kullanıcı yorumlarının duygu analizi için transformer modellerinin etkili bir şekilde uygulanabileceğini göstermiştir. Çeşitli modellerin performansları analiz edilmiş ve eğitim/çıkarım süreleri, doğruluk, karmaşıklık matrisi ve diğer sınıflandırma metrikleri gibi kriterler ışığında değerlendirilmiştir. Sonuçlar, farklı modellerin farklı avantajlar sunduğunu ve her bir modelin belirli senaryolar için uygun olduğunu ortaya koymuştur.

* 1. BERT

BERT, %88.18 doğruluk oranıyla genel olarak yüksek bir performans sergilemiştir. Modelin çift yönlü bağlam anlayışı, yorumlardaki anlam ilişkilerini güçlü bir şekilde kavramasını sağlamıştır.

Eğitim süresi ve doğruluk dengesi ile özellikle geniş veri kümelerinde duygu analizi için etkili bir seçenek olduğu kanıtlanmıştır.

* 1. RoBERTa

RoBERTa, %86.56 doğruluk oranı ile yüksek bir performans sergilemiştir. Özellikle metinlerdeki ince detayları anlamada başarılı olmuştur.

Eğitim süresi daha kısa olmasına rağmen, modelin doğruluk oranı BERT’e oldukça yakındır ve hızlı sonuç almak isteyen sistemler için uygun bir çözüm sunmaktadır.

* 1. DeBERTa

DeBERTa, %85.54 doğruluk oranı ile metinlerin bağlamını anlamada güçlü bir modeldir. Ancak, eğitim süresi diğer modellere göre daha uzun olmuştur.

Daha karmaşık yapıya sahip yorumların analizinde etkili olduğu gözlemlenmiştir.

* 1. DistilGPT-2

DistilGPT-2, %82.57 doğruluk oranıyla diğer modellere kıyasla daha düşük bir performans sergilemiştir. Bunun nedeni, modelin temel olarak metin üretimi için tasarlanmış olmasıdır.

Ancak, sınırlı kaynaklarla çalışan sistemler için daha hızlı ve hafif bir çözüm sunması nedeniyle avantajlıdır.

* 1. ALBERT

ALBERT modeli, %84.50 doğruluk oranı ile dengeli bir performans sergilemiştir. Daha az bellek kullanımı ile dikkat çekmekte ve geniş veri kümelerinde etkili bir seçenek olarak öne çıkmaktadır.

Genel Değerlendirme

Bu çalışmada kullanılan modellerin her biri, farklı avantaj ve dezavantajlara sahiptir. BERT ve RoBERTa, genel duygu analizi için güçlü performans gösterirken, DeBERTa gibi modeller daha karmaşık metinlerde etkili olmuştur. DistilGPT-2, hız ve kaynak verimliliği açısından öne çıkarken, ALBERT ise düşük bellek gereksinimi ile dikkat çekmiştir.

Sonuç olarak, farklı kullanım senaryolarına göre model seçimi yapılması gerektiği, bu çalışmanın önemli bir bulgusu olarak ortaya çıkmıştır. Elde edilen sonuçlar, transformer modellerinin e-ticaret ve forum sitelerinden elde edilen kullanıcı yorumlarının analizi için başarılı bir şekilde kullanılabileceğini göstermektedir.

##### sonuç

Bu çalışma, e-ticaret ve forum sitelerinden elde edilen kullanıcı yorumlarının duygu analizi için transformer modellerinin potansiyelini ortaya koymuştur. Çalışmada kullanılan BERT, RoBERTa, DeBERTa, ALBERT ve DistilGPT-2 modelleri, farklı metrikler ve performans ölçütleri açısından değerlendirilmiştir.

BERT modeli, doğruluk oranı ve genel performans açısından yüksek bir başarı sergileyerek kullanıcı yorumlarının analizinde etkili bir model olduğunu göstermiştir. RoBERTa ve DeBERTa, özellikle karmaşık bağlamları anlamlandırmada güçlü bir performans sergilerken, ALBERT düşük bellek kullanımı ile dikkat çekmiştir.

DistilGPT-2 modeli, metin üretimi için tasarlanmış olmasına rağmen, hızlı sonuçlar üretmesi ve kaynak verimliliği ile sınırlı sistemler için avantajlı bir seçenek olmuştur.

Elde edilen sonuçlar, farklı kullanım senaryoları için en uygun transformer modelinin belirlenmesinin önemini vurgulamaktadır. Transformer modelleri, kullanıcı yorumlarının duygu analizi görevinde etkili bir şekilde kullanılabilir ve e-ticaret, müşteri geri bildirim analizi gibi alanlarda önemli bir katkı sağlayabilir.

##### referanslar

1. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is All You Need. Advances in Neural Information Processing Systems (NeurIPS).
2. Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics (NAACL-HLT).
3. Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., ... & Stoyanov, V. (2019). RoBERTa: A Robustly Optimized BERT Pretraining Approach. arXiv preprint arXiv:1907.11692.
4. He, P., Liu, X., Gao, J., & Chen, W. (2020). DeBERTa: Decoding-enhanced BERT with Disentangled Attention. arXiv preprint arXiv:2006.03654.
5. Sanh, V., Debut, L., Chaumond, J., & Wolf, T. (2019). DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter. arXiv preprint arXiv:1910.01108.
6. Lan, Z., Chen, M., Goodman, S., Gimpel, K., Sharma, P., & Soricut, R. (2020). ALBERT: A Lite BERT for Self-supervised Learning of Language Representations. International Conference on Learning Representations (ICLR).
7. Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., & Sutskever, I. (2019). Language Models are Unsupervised Multitask Learners. OpenAI GPT-2 Technical Report.