

BERT Kullanarak Sahte Yorum Sınıflandırma Ve Performans Karşılaştırılması

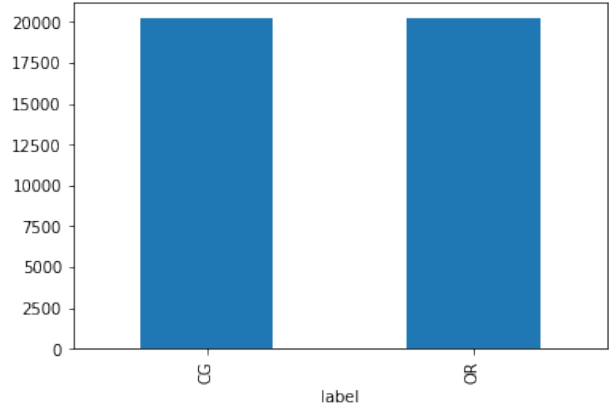
Mert AKÇAY

Ege Üniversitesi, İzmir TR

Haziran 19, 2022

1 Giriş

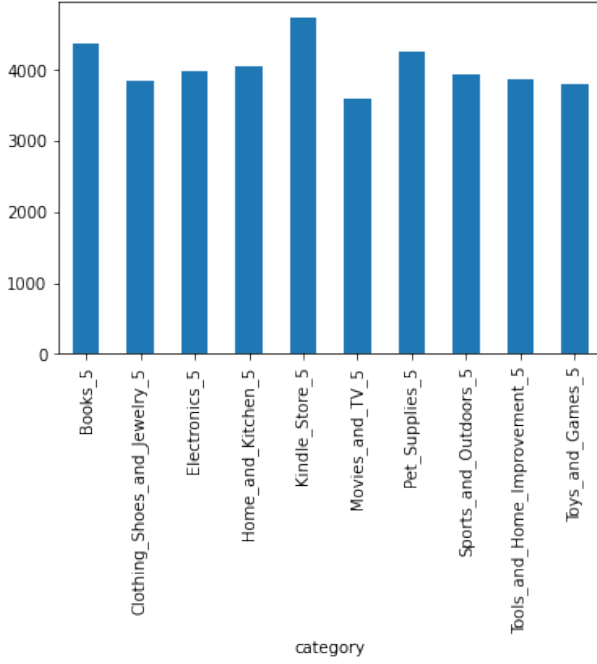
Son gelişmelerle internetin hayatımızda yaygınlaşması ve teknolojinin gelişmesi ile e-ticaret oldukça popülerdir. Kişiler seçtikleri ürünleri elinde inceleyemedikleri için çeşitli kontrol birimlerini kullanmaktadır. Bunlardan birisi ise kullanıcı yorumlarıdır. Kullanıcılar yüksek yıldız almış, fazla yorum içeren ürünleri fiyat farkı olsa bile tercih etmektedir ancak çeşitli kötü niyetli kullanıcılar tarafından bu sistem istismar edilmektedir. Bu yayında amaçlanan yaklaşım kullanıcı yorumlarını BERT modeli kullanarak gerçek veya sahte olarak sınıflandırarak BERT modelinin performans karşılaştırmalarını yapmaktır.



Verisetinin sahip olduğu kategori değerleri model eğitim aşamasında eğitim dışı edilecektir. Kategoriler de ise yorum dağılımı gibi bir ekstrem durum bulunmamaktadır.

1.1 Veri Seti

Verisetindeki gerçek ve sahte yorum dağılımı eşit olarak dağılmıştır. Minor ağırlığa sahip bir değer bulunmamaktadır. Eğitim seti boyutu: 32345, Validasyon seti boyutu: 4043, Test seti boyutu: 4044



Dataset: <https://osf.io/tyue9/>

1.2 Yorumların Özellikleri

İncelenen veriler sadece sıfat kullanımı, noktalama işaretlerinin yoğun bir şekilde atlanması ve kısa yorumlara dikkat çekmiştir.

1.3 BERT(Bidirectional Encoder Representations from Transformers)

1.3.1 BERT Çıkış Noktası

BERT(Bidirectional Encoder Representations from Transformers)' in önceki NLP yaklaşımlarından önemli artışı soldan sağa sequential izlemek yerine çift yönlü olarak kullanmasıdır. Önceki çalışmalarda bu yöntemin imkansız olacağı söylense de Google ın geliştirdiği model'de Masked LM yaklaşımı ile mümkün kılınmıştır.

Bunun yanında Computer Vision alanında bulunan transfer learning yaklaşımında burada kullanılmaktadır.

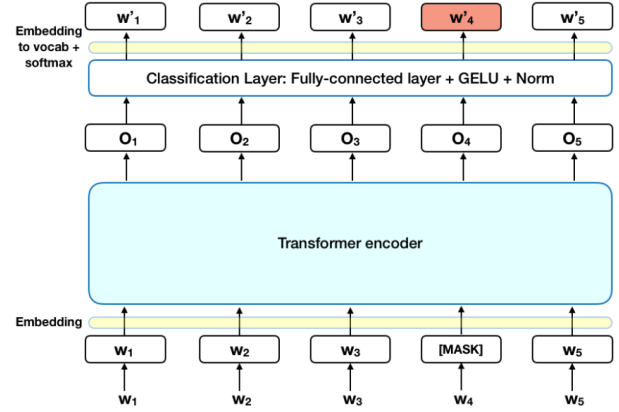
1.3.2 BERT Nedir ?

BERT kelimeler arasındaki bağlam ve duyguyu ortaya çıkarmayı transformerlar aracılığı ile ortaya çıkarmayı amaçlamaktadır. İçinde attention mechanism bulunan yapı tek bir bakışta tüm sequential yapıyı çözmeyi hedefleyerek soyut bir biçimde sanki hem soldan sağa hemde sağdan sola olarak çift yönlü okuma mekanizmasını sağlamaktadır.

Transformer'da sadece kullanılan Encoder yapısı ile bir vektör uzayı oluşturarak kullanmamızı sağlar.

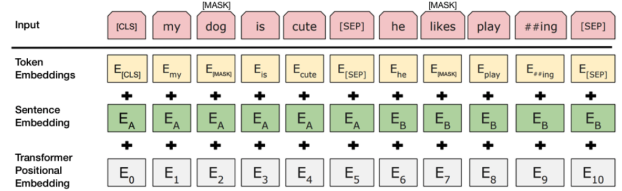
1.3.3 Masked LM (MLM)

Öğrenmeyi dengelemek amacıyla input textimizdeki verilerin %15 lik bir oranda maskelenir(boş atanır). Daha sonra bu maskelenen değerleri softmax aktivasyon fonksiyonundan geçirerek tahminlemeye çalışır. Burada tahminlenen kelimeler ile bilinen kelimeler arasındaki farklı hata fonksiyonu olarak belirterek modeli eğitir.



1.3.4 Next Sentence Prediction (NSP)

NSP BERT için 2 tane cümlecik alır. Bu cümleciklerin yerler önce veya sonra gelip gelmediğini sorgular. Eğer bu cümlecikler mantıksal olarak birbiri ile bağlantısı mümkün ise doğrudur olarak işaretleyerek eğitimi sağlar.



2 Eğitim Aşaması

2.1 Yaklaşımlar

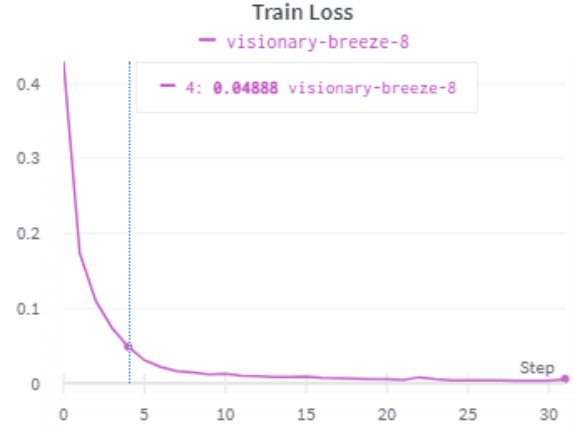
Çeşitli unsupervised NLP algoritmaları incelendi ve denendi ancak problem başarı olarak yeterli bulunmadığı için supervised yöntemleri kullanıldı. Mevcut BERT modelinin üzerine eklenen Vanilla NN ile sınıflandırması sağlandı.

2.2 Newbie Model

Öğrenmeyi en etkili bir şekilde sağlayan optimizasyonu bulmak amacıyla çeşitli optimizasyonların performansları şu şekildedir;

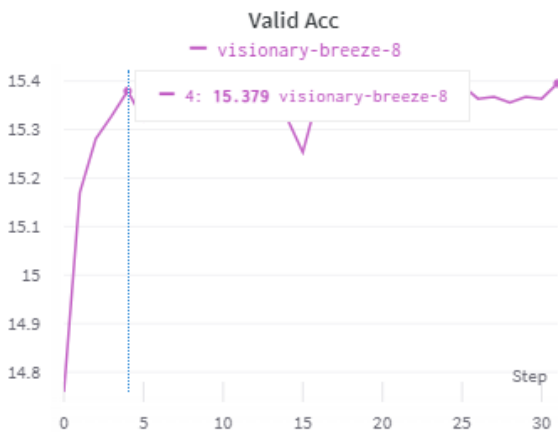
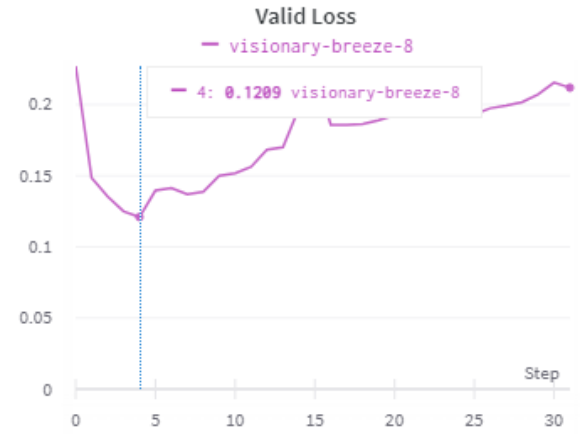
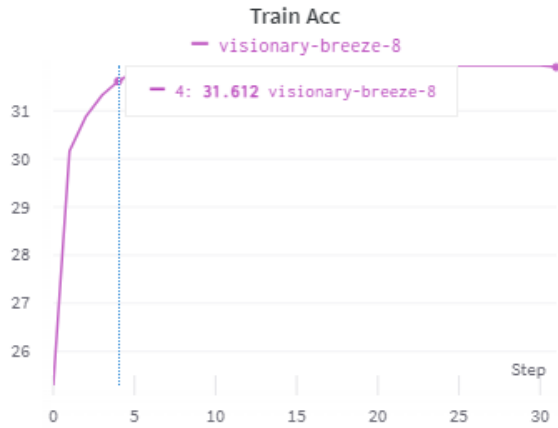
Table 1: Epoch: 32 - LR: 0,000001 - Train_Batch: 16 - Valid_Batch: 8

Optimizer	Train_Acc	Valid_Acc	Train_Loss	Valid_Loss
AdamW	0,974	0,957	0,005	0,018
Adam	0,998	0,964	0,001	0,018
RMSProp	0,978	0,957	0,014	0,028
SGD	0,973	0,95	0,016	0,03



2.3 En İyi Model

32 Epoch, 0.01-0.0000001 arasındaki yapılan LR search ile en iyi LR olarak 0.000001 eğitilen modelde 4. Epochtan sonra eğitim verisini ezberlemeye başladığı gözlemdi. Validasyon ve Eğitim verileri benzer dağılım göstermesine rağmen bu durum gözlenmiştir. İkili sınıflandırma yapıldığı için Eğitim verisinde %98,7875 başarı oranı Validasyon veri setinde ise %96,11875 başarı oranı yakalanmıştır.



3 Kaynakça

- [1][https://en.wikipedia.org/wiki/BERT_\(language_model\)](https://en.wikipedia.org/wiki/BERT_(language_model))
- [2]<https://huggingface.co/blog/bert-101>
- [3]<https://huggingface.co/models?language=en&sort=downloads>
- [4]<https://towardsdatascience.com/bert-explained-state-of-the-art-language-model-for-nlp-f8b21a9b6270>
- [5]<https://arxiv.org/pdf/2010.15296.pdf>
- [6]<https://arxiv.org/pdf/1810.04805.pdf>
- [7]<https://www.section.io/engineering-education/classification-model-using-bert-and-tensorflow>
- [8]<https://arxiv.org/abs/1905.05583>
- [9]<https://arxiv.org/abs/2111.15379>
- [10]<https://arxiv.org/pdf/2202.00795.pdf>