



Doğal Dil İşlemeye Kavramsal Bir Bakış Proje Raporu

24501097

Yusuf Enes Kurt

enes.kurt1@std.yildiz.edu.tr

Ders Yürütücsü

Prof. Dr. Banu Diri

23.06.2025

Proje Nedir?

Bu proje, 2022 Dünya Kupası ile ilgili atılan tweet'ler üzerinden İngilizce duyu analizi (sentiment analysis) yapmayı hedeflemektedir. Tweet'ler pozitif, negatif ve nötr olmak üzere üç sınıfa ayrılmıştır. Proje kapsamında dört farklı yöntemle duyu analizi gerçekleştirilmiştir:

1. **Elle Yazılmış Basit Bir Sinir Ağı:** Temel bir yapay sinir ağı modeli sıfırdan oluşturularak sınıflandırma (classification) işlemi yapılmıştır.
2. **Hazır Modelin İncelenmesi ve İnce Ayar (Fine-tuning):** Hugging Face üzerinden seçilen güçlü bir önceden eğitilmiş model, proje verisiyle yeniden eğitilerek sınıflandırma gerçekleştirilmiştir.
3. **Duyu Yoğunluğu Tahmini (Regression):** Aynı hazır model bu kez sınıflandırma yerine -1 ile 1 arasında duyu yoğunluğu skorlarını tahmin edecek şekilde regresyon problemi olarak ele alınmıştır.
4. **Skorlardan Sınıfa Dönüşümü:** Regresyon çıktıları belirli eşiklerle sınıflandırmaya çevrilmiş ve bu yöntem ile önceki sınıflandırma modellerinin doğruluk, F1 skoru gibi metriklerle karşılaştırılması yapılmıştır.

Bu çok yönlü çalışma ile hem temel hem ileri düzey yöntemlerin duyu analizindeki performansı incelenmiş, sınıflandırma ve regresyon tabanlı yaklaşımların kıyaslaması yapılmıştır.

Projenin Katkı Sağlayabileceği Problemler

1. Büyük turnuvalar sırasında halkın duyu durumunu izleyerek organizasyonlara yönelik genel memnuniyet ölçülebilir.
2. Takım veya sponsor markaların sosyal medya üzerindeki algısı izlenebilir, kriz durumları erken tespit edilebilir.
3. Regresyon çıktıları sayesinde daha hassas filtreleme ve sıralama algoritmaları geliştirilebilir (örneğin haber veya yorum tavsiyesi).

Dataset

Bu projede kullanılan veri seti, 2022 FIFA Dünya Kupası süresince sosyal medya platformu Twitter'da paylaşılan tweet'lerden oluşmaktadır. Tweet'ler, içeriklerine göre üç sınıfa ayrılmıştır: **pozitif, nötr ve negatif**.

Toplam veri sayısı: 22.524 tweet

Label Türleri: Index, Date Created, Number of Likes, Source of Tweet, Tweet, Sentiment

positive: 8489 tweets (37.69%)
neutral: 8251 tweets (36.63%)
negative: 5784 tweets (25.68%)

Yapılan İşlemler

1. Gereksiz kolonlar temizlenmiş, Yalnızca Tweet ve Sentiment sütunları tutulmuştur.
2. Her sınıfın eşit sayıda (5784) tweet rastgele örneklenerek dengeli bir veri seti oluşturulmuştur.
3. %70 eğitim, %15 doğrulama ve %15 test seti olmak üzere veri seti üçe ayrılmıştır. Her sette sınıf oranları korunmuştur.
4. “sentence-transformers/all-MiniLM-L6-v2” hazır modeliyle tweet’ler vektör uzayına gömülüdür. Bu işlem sadece 1. Model için kullanılmıştır.

Sentiment		Tweet	Sentiment
negative	5784	0 How in the name of Jesus was that not a good g...	negative
neutral	5784	1 @JioCinema your app is lagging on smart tv. Pl...	negative
positive	5784	2 Hey @benske31 we gotta watch some @CanadaSo...	neutral
		3 Watching Qatar play is like watching Sunday le...	neutral
		4 Are you guys broadcasting in the kitchen as al...	neutral

```
{'train': {'embeddings': array([[-0.05856002,  0.03054745, -0.004129 , ..., -0.01714644,
   -0.0281125 ,  0.0377028 ],
 [ 0.00328097, -0.04143158,  0.01772959, ..., -0.06149257,
  0.02362919, -0.05104531],
 [ 0.02373547,  0.04418331,  0.08109222, ..., -0.03454321,
  -0.0902932 , -0.00294139],
 ...,
 [ 0.00969438,  0.06337392,  0.07588311, ..., -0.04241866,
  -0.1101367 ,  0.00172768],
 [ 0.00117593,  0.03903282,  0.0147438 , ..., -0.06552655,
  0.01816062, -0.02369412],
 [-0.01741014,  0.05127728,  0.00107904, ..., -0.04754308,
  0.02919964,  0.02829987]], dtype=float32),
```

Embedding Modeli

sentence-transformers/all-MiniLM-L6-v2, kısa metinler için hızlı ve etkili vektör temsilleri (embedding) üreten hafif bir dil modelidir. 6 katmanlı MiniLM mimarisi üzerine kuruludur ve özellikle duygusal analizi ve benzerlik karşılaştırma gibi doğal dil işleme görevlerinde yüksek performans gösterir. Düşük işlem gücü gerektirir.

Basit Bir Yapay Sinir Ağı ve Sonuçlar

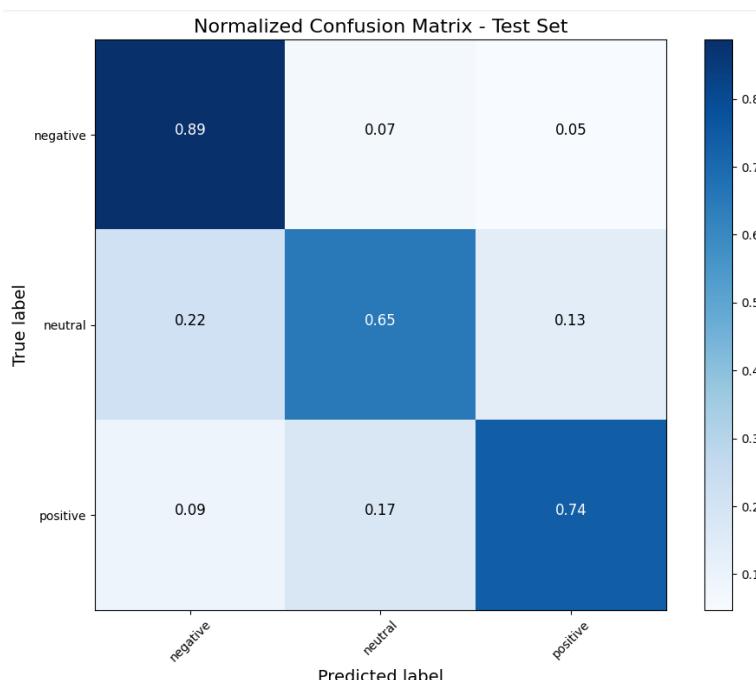
Önceden sentence-transformers/all-MiniLM-L6-v2 modeliyle vektörleştirilen tweet embedding'leri kullanılarak Keras ile basit bir yapay sinir ağının kurulmuş olması.

Layer (type)	Output Shape	Param #
dense (Dense)	(None, 128)	49,280
dropout (Dropout)	(None, 128)	0
dense_1 (Dense)	(None, 64)	8,256
dropout_1 (Dropout)	(None, 64)	0
dense_2 (Dense)	(None, 32)	2,080
dropout_2 (Dropout)	(None, 32)	0
dense_3 (Dense)	(None, 3)	99

Genel Metrikler

- **Test Accuracy:** 0.7580
- **Precision (Macro):** 0.76
- **Recall (Macro):** 0.76
- **F1-Score (Macro):** 0.76

Normalize Edilmiş Confusion Matrisi



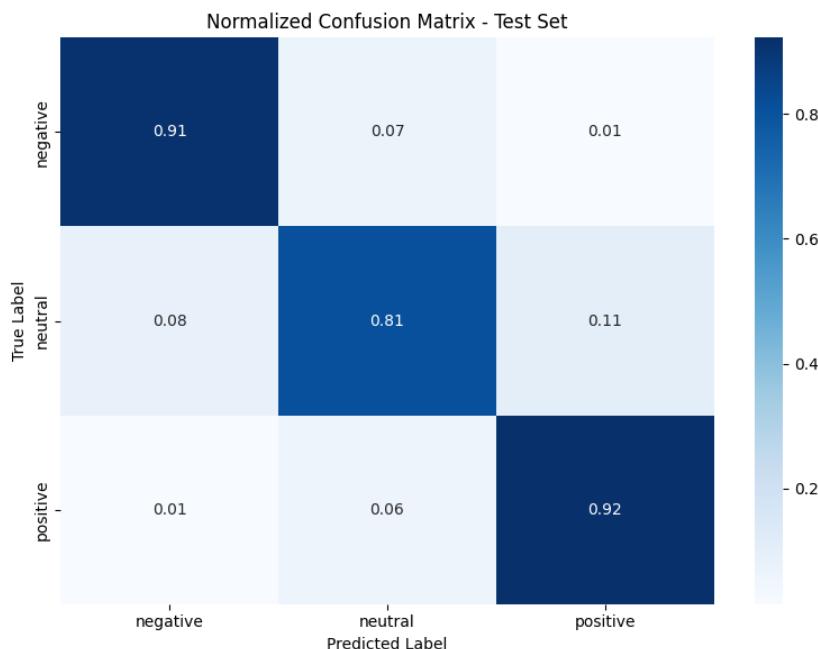
Finetuned Text Classification Model ve Sonuçlar

Kullanılan model, Hugging Face üzerinden sağlanan **cardiffnlp/twitter-roberta-base-sentiment-latest** adlı hazır bir duygusal analizi modelidir. Bu model, Twitter verileri üzerinde eğitilmiş RoBERTa tabanlı bir dil modelidir ve tweet'leri üç sınıfa (**pozitif, nötr, negatif**) ayırmak için optimize edilmiştir. Sosyal medya diline özel tasarlandığı için kısa ve gündelik ifadelerde yüksek performans göstermektedir.

Genel Metrikler

- **Test Accuracy:** 0.8813
- **Precision (Macro):** 0.8807
- **Recall (Macro):** 0.8813
- **F1-Score (Macro):** 0.8806

Normalize Edilmiş Confusion Matrisi



Finetuned Regression Model ve Sonuçlar

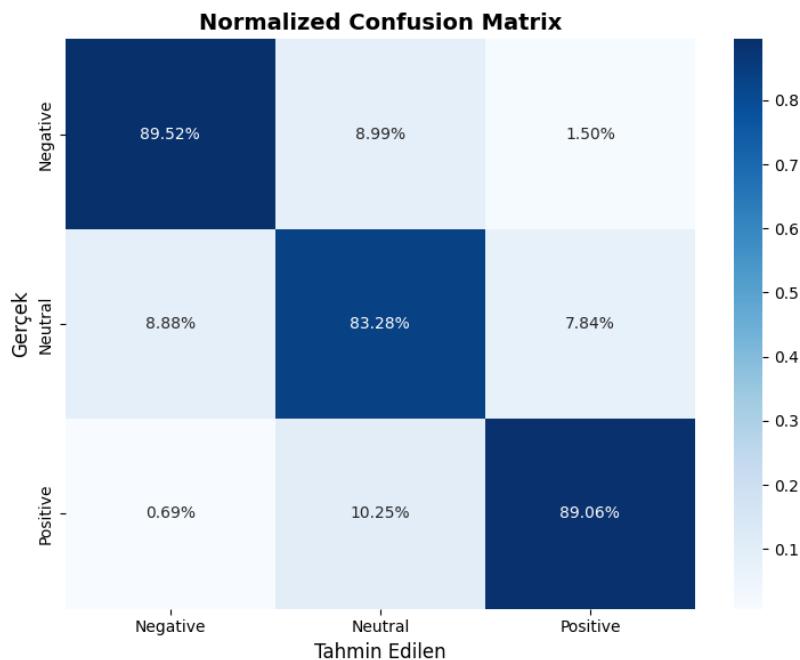
Bu model aynı hazır modelinin üstüne özel olarak geliştirilmiş bir **regresyon başlığı** eklenerek tweet'lerin duyu yoğunluğunu sürekli bir değer olarak tahmin etmek üzere uyarlanmıştır. Modelin çıktısı -1 ile 1 arasında olup bu aralıkta **-1 en negatif, 0 nötr, 1 ise en pozitif** duyguyu temsil eder.

Daha sonrasında en iyi eşik değerler bulunarak regresyon, sınıflandırmaya dönüştürülmüştür.

Genel Metrikler

- **Test Accuracy:** 0.8728
- **Precision (Macro):** 0.8736
- **Recall (Macro):** 0.8728
- **F1-Score (Macro):** 0.8732

Normalize Edilmiş Confusion Matrisi

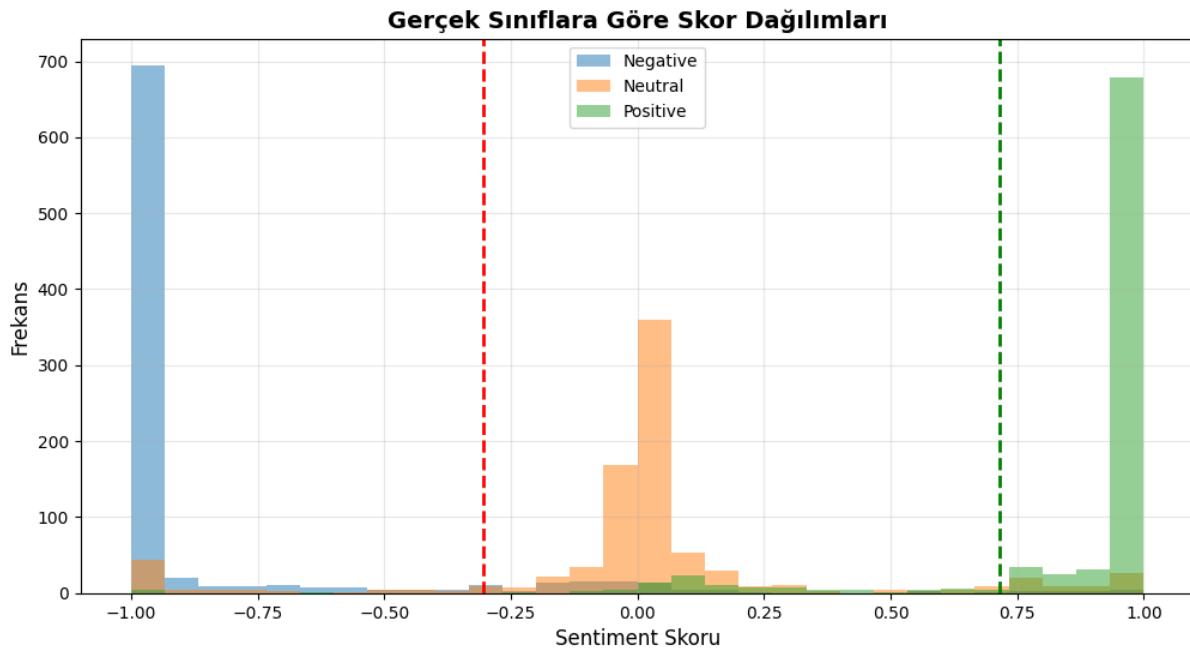


Optimal Eşikler

Negatif eşik: -0.305

Pozitif eşik: 0.715

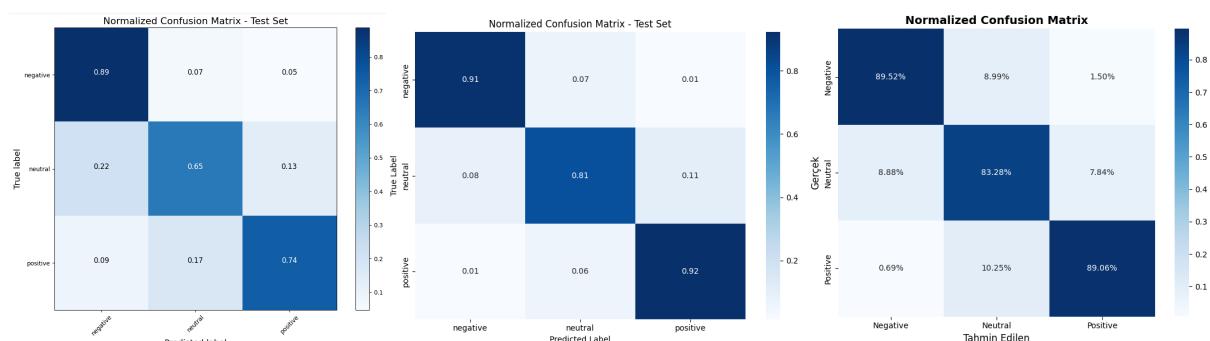
En iyi F1 skoru: 0.8732



Model Performans Karşılaştırması

Model Adı	Test Doğruluğu (Accuracy)	Precision (Makro)	Recall (Makro)	F1-Skoru (Makro)
Basit Yapay Sinir Ağısı	0.7580	0.7600	0.7600	0.7600
Finetuned Sınıflandırma Modeli	0.8813	0.8807	0.8813	0.8806
Finetuned Regresyon Modeli	0.8728	0.8736	0.8728	0.8732

Confusion Matris Karşılaştırması



Genel Değerlendirme

- Basit sinir ağı, düşük işlem gücüyle temel seviye başarı sunar ancak dilin bağlamını yakalamada daha zayıftır. Negatif bulmada diğer modellere yakın başarılar elde etmiştir.
- En yüksek doğruluk, doğrudan sınıflandırma yapan finetuned RoBERTa modelinde elde edilmiştir. Pozitif ve negatif bulmada oldukça başarılıdır.
- Regresyon modeli, sınıflandırmaya dönüştürülmüşe rağmen ona oldukça yakın sonuçlar üretmiştir ve ek analiz imkânı sunmuştur. En iyi eşik değerler uygulandığında nötr ifadeleri bulmada en başarılı model olmuştur.