

# Türkçe Şarkı Sözleriyle Müzik Türü Sınıflandırması

Hasan Erdem AK Oğuz Öçal

Bilgisayar Mühendisliği Bölümü

TOBB Ekonomi ve Teknoloji Üniversitesi, Ankara, Türkiye

hak@etu.edu.tr

oocal@etu.edu.tr

## Özet

Bu çalışmada, Doğal Dil İşleme tekniklerini kullanarak Türkçe şarkı türlerini, sözlerine göre sınıflandırmayı amaçlayan bir proje gerçekleştirdik. Oluşturduğumuz şarkı sözü külliyatını, makine öğrenimi modelleri ve Bi-LSTM ile eğittik. En yüksek başarımla, 0.77 ile Random Forest ve Bi-LSTM ile elde edildi. Projeden elde edilen sonuçlar, Türkçe müzik endüstrisine pazar trendlerini ve tüketici tercihlerini anlamada yardımcı olma potansiyeline sahiptir ve Türkçe müzik dinleyicilerinin yeni müzikler keşfetmelerine katkıda bulunabilir.

## 1 Giriş

Müzik; düşünceleri, duyguları ve hisleri aktaran evrensel bir iletişim aracıdır. Gelişen teknoloji sayesinde müziğe erişim ve müzik üretme imkânları artmış, bu da müziği hayatımızın vazgeçilmez bir parçası haline getirmiştir. Müzik türlerine göre sınıflandırma, dinleyicilerin zevklerine ve tercihlerine uygun yeni müzikler keşfetmeleri açısından önemlidir. Günümüzde, müzik endüstrisine önemli yatırımlar yapan şirketler, çeşitli müzik öneri sistemleri kullanarak uygulamalarını geliştirmeye çalışmaktadır. Bu öneri sistemlerinin temel araçlarından biri yapay zeka yöntemleridir.

Müzik; melodi, ritim ve söz gibi öğelerden oluşur. Melodi ve ritim çeşitli alet ve enstrümanlarla yapılırken, sözler dil ile söylenir ve yazı ile ifade edilebilir. Farklı müzik türleri genellikle melodi ve ritme göre ayrılırken, şarkı sözlerinin de bu sınıflandırmada önemli bir rolü vardır. Şarkı sözlerinin kendine özgü yapısal özellikleri, hem biçimsel hem istatistiksel özniteliklerin elde edilmesine elverişlidir. Yapay zeka yöntemlerinin temelini oluşturan doğal dil işleme teknikleri kullanılarak sözlere göre de şarkıları sınıflandırmak mümkündür.

Bu çalışmada, Doğal Dil İşleme (DDİ) tekniklerini kullanarak Türkçe şarkı türlerini, söz-

lerine göre sınıflandırmayı amaçladık. Kullanılan şarkı türleri; halk müziği, sanat müziği, ilahi, pop, rock ve rap'tir. Geniş bir şarkı sözü külliyatı oluşturuldu, elde edilen ham veriler işlendikten sonra makine öğrenimi modelleri ve Bi-LSTM ile eğitildi. Performans, standart ölçütler kullanılarak değerlendirildi.

Bu projenin önemi, Türkçe müzik endüstrisine pazar trendlerini ve tüketici tercihlerini anlamada yardımcı olma potansiyelinde yatmaktadır. Doğru tür sınıflandırması ile Türkçe müzik endüstrisi, daha hedefli pazarlama stratejileri ve promosyonlar oluşturabilir. Ayrıca, projenin sonuçları, Türkçe müzik dinleyicilerinin yeni müzikler keşfetmelerine yardımcı olabilir ve müzik endüstrisi üzerinde olumlu bir etkiye sahip olabilir. Bu çalışma, DDİ ve Türkçe müzik sınıflandırması alanında daha fazla araştırma için bir referans olarak hizmet edebilir ve bu alanda ilerlemelere yol açabilir. Ayrıca, dilbilim, müzikoloji ve bilgisayar bilimi alanlarının birleşimiyle, Türkçe müziğe özgü yapısal özellikler ve sınıflandırma yöntemlerinin daha derin bir anlayışını sağlar.

DDİ tekniklerinin Türkçe müzik türlerini sınıflandırmada kullanılması, dilbilim ve kültür açısından zengin olan Türkçe'nin daha iyi anlaşılmasına ve değerlendirilmesine katkıda bulunabilir. Bu yaklaşım, kültürel farklılıkları ve benzerlikleri daha iyi anlamak, dil ve müzik arasındaki ilişkiyi keşfetmek ve Türkçe şarkı sözlerindeki duygusal ve anlamsal unsurları analiz etmek için yeni fırsatlar sunar.

Sonuç olarak, bu çalışma, Türkçe müzik endüstrisine pazar trendleri ve tüketici tercihleri konusunda daha iyi bilgi sağlayarak, sektördeki şirketlerin daha etkili pazarlama stratejileri geliştirmesine ve müzik dinleyicilerinin zevklerine uygun yeni müzikler keşfetmesine yardımcı olabilir. Ayrıca, bu çalışma DDİ ve Türkçe müzik sınıflandırması alanında daha fazla araştırmaya il-

ham verebilir ve bu alanlardaki bilgi birikimine katkıda bulunabilir.

## 2 İlgili Çalışmalar

### 2.1 Veri Seti ve Ön işleme Yaklaşımları

Veri seti oluşturmak için şarkı sözleri Kaggle<sup>1</sup>, LyricsFly<sup>2</sup>, LyricsFreak<sup>3</sup>, LyricWiki<sup>4</sup> gibi kaynaklardan alınmıştır. Bazı çalışmalarda (Hızlısoy and Tüfekci, 2021) ise ilgili dillerde veri setleri olmamasından dolayı veri setleri, çalışmaları gerçekleştirenler tarafından oluşturulmuştur.

Bu projede, ilgili çalışmalarda olduğu gibi öznitelik elde etmede kullanılan modellerden önce farklı ön işleme teknikleri kullanılmıştır. İlgili makalelerde, metin ilk olarak normalleştirilip paragraflara bölünmüştür (Mahedero et al., 2005; Çoban, 2017; Fell and Sporleder, 2014; Logan et al., 2004; Mikolov et al., 2013). Tüm modellerde ortak olarak metin içeriği üzerinde küçük harf dönüşümü uygulanmış ve öznitelik olarak kabul edilmeyen bilgiler ve tüm karakterler temizlenmiştir. Ayrıca kıtalar arasında tek satır boşluğu olması sağlanmıştır. N-Gram ve BoW (Bag of Words) modelde ortak işlemlere ek olarak harfler dışında tüm noktalama işaretleri ve karakterler de metin içeriğinden çıkarılmış (Boonyanit and Dahl, 2022; Laurier et al., 2008), terimler için uzunluk ve frekans filtresi uygulanmamıştır.

### 2.2 Temel Modeller

Genel çalışmaları incelediğimizde temel alınan modeller Lojistik Regresyon (Boonyanit and Dahl, 2022; Tsaptsinos, 2017; Yang and Lee, 2009), uygulamanın rastgele bir seçim yapması durumunda istatistiksel başarımın ne olabileceği, temel SVM algoritması, LSTM (Tsaptsinos, 2017) gibi modeller seçilmiştir. Sadece ses özelliklerini kullanan bir sınıflandırıcıyı temel model olarak kabul eden çalışmalar da vardır (Mayer and Rauber, 2011; McKay et al., 2010).

### 2.3 Makine Öğrenmesi Tabanlı Yaklaşımlar

Makine Öğrenimi modellerini kullanan çalışmalarda çeşitli öznitelikler kullanılmıştır. İlgili bir makalede (Çoban and Karabey Aksakalli, 2017),

anlamsal işlemler için Zemberek<sup>5</sup> kullanılmıştır. Metin içerisindeki durak kelime oranının bulunması aşamasında ise Lucene API'de<sup>6</sup> bulunan Türkçe durak kelimeler listesi kullanılmıştır. Şarkı sözlerindeki kelimelerin ve kelime gruplarının sıklığı da sınıflandırma için kullanılan öznitelikler arasındadır (Ying et al., 2012).

Başka bir çalışmada ise metin sınıflandırma problemi için bir dilin tüm kelimelerini, bu kelimelerin tanımlarını ve aralarındaki anlambilimsel ilişkileri içeren bir veritabanı olan WordNet<sup>7</sup> kullanılarak kelimelerin üst başlıkları yani kendisiyle ilgili olan ama kendinden daha genel bir anlama sahip olan kelime öznitelikleri kullanarak kelime çuvalı gösterimi, üst isim yoğunluğu olarak değiştirilerek oluşturulan özniteliklerin sağladığı anlamsal bilgi kullanılmıştır (Scott and Matwin, 1998). Bu çalışmalardan yola çıkarak biz de çalışmamızda kelimelerle ilgili çeşitli öznitelikleri (ortalama kelime sayısı, kelime tekrarı, kelime türü oranları vs.) ve duygu analizi sonuçlarını veri setimize ekledik.

Svm algoritması, Karar ağaçları gibi anlamlandırılabilmesi yapay sinir ağlarına göre daha kolay olan algoritmalar sıklıkla tercih edilen ve başarılı sonuçlar veren algoritmalar (Ying et al., 2012; Mayer et al., 2008). Birçok farklı model ile yürütülen bir çalışmada, sınıflandırma modellerinde deneyler 10-kat çapraz doğrulama yöntemi ile gerçekleştirilmiştir (Mahedero et al., 2005). Bazı çalışmalarda ise Makine Öğrenimi metodlarından olan ve Doğal Dil İşleme'de sıkça kullanılan Naive Bayes, K-en yakın komşu, rastgele orman modelleri kullanılmıştır (Mahedero et al., 2005; Kleedorfer et al., 2008; Fell and Sporleder, 2014; Laurier et al., 2008). Ayrıca, topluluk (ensemble) sınıflandırıcıyı kullananlar da bulunmaktadır (Mayer and Rauber, 2011). Bu projede de farklı makine öğrenimi modelleri kullanılmış ve birbirleriyle karşılaştırılmıştır. Bunlar, Random Forest, SVM, Naive Bayes ve XGBoost'tur.

### 2.4 Derin Öğrenme Tabanlı Yaklaşımlar

Literatürdeki çalışmaların birçoğunda, kullanılan veri setindeki her bir şarkının temsilinde, metin sınıflandırma çalışmalarında yaygın olarak kullanılan kelime gömmeleri kullanılmıştır (Boonyanit and Dahl, 2022; Kumar et al., 2018;

<sup>1</sup><https://www.kaggle.com/>

<sup>2</sup><http://www.pearltrees.com/u/1169110-lyricsfly-lyrics-database>

<sup>3</sup><https://www.lyricsfreak.com/>

<sup>4</sup><https://lyricwiki.wordpress.com/>

<sup>5</sup><https://github.com/ahmetaa/zemberek-nlp>

<sup>6</sup>[https://lucene.apache.org/core/8\\_0\\_0/core/overview-summary.html](https://lucene.apache.org/core/8_0_0/core/overview-summary.html)

<sup>7</sup><https://wordnet.princeton.edu/>

Tsaptinos, 2017; Çoban, 2017). Kelime gömme yöntemlerinden ise Word2Vec ve GloVe gömmeleri kullanılmıştır (Kumar et al., 2018; Boonyanit and Dahl, 2022; Tsaptinos, 2017; Çoban, 2017). Veri gösterimi olarak döküman vektörleri(doc2vec) de kullanılmıştır (Çoban and Karabey Aksakalli, 2017). Kelime gömme yöntemini kullanan projelerden biri, Word2Vec'in 2 farklı versiyonunu kullanmış. Bunlar, CBOW (Continuous Bag of Words (Sürekli Kelime Çuvalı)) Word2Vec ve TF-IDF (Term Frequency - Inverse Document Frequency (Terim Sıklığı-Ters Belge Sıklığı)) ile birlikte kullanılan Word2Vec'tir (Kumar et al., 2018; Mayer and Rauber, 2011; Fell and Sporleder, 2014; Laurier et al., 2008). Ayrıca çalışmada elde edilen sonuçlara göre TF-IDF ile birlikte Word2Vec yöntemi kullanılarak elde edilen vektörel temsilde daha iyi bir başarımla elde edildiği belirtilmiştir (Kumar et al., 2018).

Bu bilgilerden yola çıkarak biz de kullandığımız derin öğrenme modelinde, kelime gömme yöntemini kullandık. Kelime gömme modeli olarak ise Türkçe metinler üzerinde eğitilmiş olan bir kelime gömme modeline sahip olan Fast-Text<sup>8</sup> kelime gömmelerini kullandık.

## 2.5 Türkçe Şarkı Türü Sınıflandırma Hakkında Yapılan Çalışmalar

Literatürde, Türkçe şarkı sözlerini kullanarak şarkı türü sınıflandırması yapan az sayıda çalışma bulunmaktadır (Çoban and Özyer, 2016; Çoban, 2017; Çoban and Karabey Aksakalli, 2017; Hızlısoy and Tüfekci, 2021). Bu çalışmalarda kullanılan veri setleri, Türkçe şarkı sözlerinden oluşan hazır bir veri tabanı olmadığı için çalışmayı gerçekleştirenler tarafından oluşturulmuştur (Çoban and Özyer, 2016; Çoban, 2017; Çoban and Karabey Aksakalli, 2017; Hızlısoy and Tüfekci, 2021). Aynı sebepten ötürü biz de şarkı sözleri veri setimizi, çeşitli sitelerden çektiğimiz şarkı sözleriyle oluşturduk.

Yapay sinir ağları üzerine yapılan bir çalışmada, sınıflandırmayı yapmak için bu alanda daha önce kullanılmamış derin öğrenmeye dayalı Evrişimli Uzun Kısa süreli bellek derin sinir ağı (CLDNN) adı verilen bir mimari kullanılmıştır. Önerilen mimarinin ve yaygın olarak kullanılan makine öğrenme metotlarının sınıflandırma performansı oluşturulan Türkçe Müzik Veri Tabanı'nda değerlendirilmiştir (Hızlısoy and Tüfekci, 2021). Bu

<sup>8</sup><https://fasttext.cc/docs/en/crawl-vectors.html>

makalede, kullanılan modeller arasında en iyi sonucu verenin CLDNN mimarisi olduğu belirtilmiş ve en iyi sonuçlara %99,5 ile bu yöntemle ulaşılmıştır (Hızlısoy and Tüfekci, 2021).

## 3 Metodoloji/Önerilen Çözüm

### 3.1 Problem

Bu çalışmada girdi olarak verilen şarkı sözlerinin anlamsal, biçimsel veya diğer farklı öznitelikleri ile hangi müzik türüne ait olduğunun belirlenmesi problemidir. Örnek olarak şarkı sözüne karşılık elde edilmesi gereken şarkı türü aşağıdaki tablodaki gibidir.

Şarkı Sözleri	Şarkı Türü
"...Zehretme hayâtı bana cânânım." Elemlerle doldu benim her anım.	Sanat
"...Beni benimle bırak" Beni benimle bu cehennemde.	Rock

Table 1: Şarkı Sözleri ve Türleri

Çalışmada 6 farklı Türkçe müzik türü tahmin edilecektir. Bu türler: Rock, Rap, Pop, Türk Sanat Müziği, Türk Halk Müziği ve Tasavvuf müziğidir.

### 3.2 Çözüme Genel Bakış

Öncelikle, Türkçe müzik sözlerini içeren hazır bir veri seti bulunmadığından veri seti bizim tarafımızdan oluşturulacaktır. Bu veri setinde her bir şarkı türünden 200'er tane şarkı sözü, sanatçılarıyla birlikte veri setine eklenecektir. Sonrasında veriler ön işleme aşamasından geçirilerek analiz ve veri temsil modelleri için hazır hale getirilecektir.

Bu aşamada elde edilen ön işlenmiş veriler üzerinde birtakım analizler yapılacaktır: Öncelikle EDA (Exploratory Data Analysis, Keşifsel Veri Analizi) gerçekleştirilecek ardından ise her tür için, şarkı sözlerinden kelime frekansları, kelime uzunlukları ortalamaları, tekrar eden kelimelerin tekrar etme oranları ve ortalamaları, eğitilmiş modeller ile şarkıya ait duygu analizi sonuçları çıkarılacaktır. Daha sonra ise her bir müzik türlerine ait yukarıda bahsettiğimiz tüm özniteliklerin tür ortalamaları kıyaslamaları histogramlar ile gösterilecektir.

Planladığımız çözümde, verileri temsil etmek için daha önceki çalışmalarda daha iyi sonuç verdiği ortaya çıkan Fasttext kelime gömmeleri kullanılacaktır. Ayrıca bizim gözümüzde şarkıları sözlerine göre ayırt edilebilir kılan bazı yapısal ve istatistiksel öznitelikleri de eğitimde kullanmak istediğimiz için SSTF (Structural and Statistical

Text Features, Yapısal ve İstatistiksel Metin Öznitelikler) modeli de kullanılacaktır. Elde edilen sayısal vektörler ve yapısal ve istatistiksel öznitelikler Bi-LSTM (Long Short Term Memory) modeline farklı girdi katmanlarında verilerek sınıflandırma yapılacaktır. Bu çalışmada Bi-LSTM modelinin kullanılma nedeni, kelime dizilerinin sıralı yapısını dikkate alması ve kelime dizileri arasındaki bağıntıları modelleyebilmesidir. Bu nedenle, daha önceki çalışmalarda diğer modellere kıyasla daha yüksek doğruluk oranlarına ulaşmıştır (Boonyanit and Dahl, 2022; Tsaptsinos, 2017). Ayrıca Bi-LSTM'e rakip olan HAN (Hierarchical Attention Network) çok sayıda tür içeren sınıflandırmalarda daha iyi sonuç elde etse de bizim çalışmamızdaki gibi az sayıda tür içeren sınıflandırmalarda Bi-LSTM daha iyi sonuç vermiştir. Bu çalışmadaki varsayımlarımız şunlardır:

- Veri setindeki parçaların ait olduğu türler doğrudur.
- Şarkının sözleri değiştirilmemiştir.

Diğer çalışmalardan farklı olarak yapacaklarımız şunlardır: veri analizleri sonucu türler arası fark yarattığını gözlemlediğimiz öznitelikler, makine öğrenimi modelinde ve kelime gömme yöntemiyle elde edilen veri temsilleri de Bi-LSTM modelinde eğitilerek her iki modelin çıktıları birleştirilmiştir. Çalışmamızda önerdiğimiz çözümün, modelin başarısını artıracaklarını düşünüyoruz. Ayrıca özniteliklerimize ekleyeceğimiz Duygu Analizi sonuçları da tür sınıflandırmada başarımda artış sağlayacaktır.

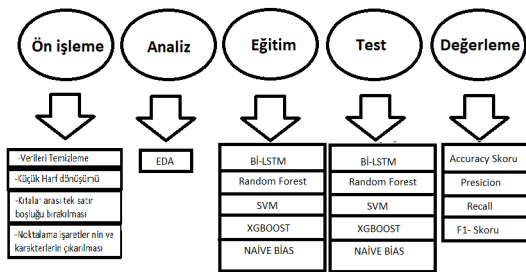


Figure 1: Model Mimarisi

### 3.3 Detaylar

Veriler ön işleme aşamasından geçirilirken aşağıdaki işlemler gerçekleştirilecektir:

- Verileri temizleme: (Satır ve satır sonlarındaki boşlukların kaldırılması, noktalama işaretlerinin

kaldırılması ve durma kelimelerinin kaldırılması)  
 - Metin içeriği üzerinde küçük harf dönüşümü,  
 - Dize sonlarının [SEP] simgesi ile işaretlenmesi,  
 - Metnin başının [CLS] simgesi ile işaretlenmesi ve BERT kelime gömme modelinin uygulanabilmesi için gerekli olan diğer işlemler gerçekleştirilecektir.

Önişlenmiş veriler, birtakım analiz aşamalarından geçirilecektir. EDA (Exploratory Data Analysis, Keşifsel Veri Analizi) gerçekleştirilecek ardından ise her tür için, şarkı sözlerinden kelime frekansları, kelime uzunlukları ortalamaları, tekrar eden kelimelerin tekrar etme oranları ve ortalamaları, eğitilmiş modeller ile şarkıya ait duygu analizi sonuçları çıkarılacaktır. Daha sonra ise her bir müzik türlerine ait yukarıda bahsettiğimiz tüm özniteliklerin tür ortalamaları kıyaslamaları histogramlar ile gösterilecektir. Eski Türkçe kelime oranının bulunma aşamasında bir kelimenin çekimli formları bir araya getirilecektir (lemmatization). Makine öğrenimi modele aşağıdaki öznitelik verileri de eklenecektir:

- Dizelerdeki cümle tekrarlarının sayısı
- Farklı bir algoritmayla elde edilen, dizelerdeki duygu analizi sonuçları
- Dizelerdeki ortalama kelime sayısı
- Dizelerdeki ortalama kelime uzunluğu
- Şarkıyı seslendiren sanatçı
- Eski Türkçe kelime oranı
- Müzik parçasındaki fiil oranı
- Müzik parçasındaki sıfat oranı
- Müzik parçasındaki zarf oranı
- Müzik parçasındaki bağlaç oranı
- Müzik parçasındaki ünlem oranı
- Müzik parçasındaki soru oranı
- Müzik parçasındaki belirlenemeyen cümle ögesi oranı

Ayrıca Türkçe'nin zengin kelime yapısı dikkate alındığında müzik sözlerindeki kelimelerin, halk arasında kullanılma dönemleri farklı olabilir. Bu yüzden, Türkçe kelime külliyatımızı Eski Türkçe ve Yeni Türkçe kelimeler olarak ayıracağız. Özellikle TSM (Türk Sanat Müziği) türünde diğer türlerden ayırt edici olarak Eski Türkçe kelimeler



daha fazla geçmektedir. Bu sebeple özniteliklerimizde, “Metinde geçen eski Türkçe kelime oranı” ve “Metinde geçen yeni Türkçe kelime oranı” özniteliklerini de ekleyeceğiz. Bu sayede TSM türünü diğer türlerden ayırt etmek için önemli bir öznitelik elde etmiş olacağız.

Daha sonra öznitelik olarak kullanacağımız şarkı sözlerinin duygu özellikleri için metin duygu analizi modeli kullanarak metnin içindeki kelimelerin anlamlarını ve ilişkilerini analiz ederek, metnin hangi duygu temasını ifade ettiğini belirleyeceğiz. Mutluluk, Sevgi, Şaşkınlık, Korku, Öfke, Üzüntü duygu temaları için önceden hazırlanmış olan TextBlob-TR<sup>9</sup> kütüphanesinde yer alan Emotion Detection modeli bizim problemimiz için metinleri duygularına göre sınıflandıracak olan model olarak seçilmiştir.

Verilerin ön işleme ve analiz aşamalarının ardından SSTF grubu özniteliklere ek olarak bizim ekleyeceğimiz özniteliklerin çıkarımı için birtakım işlemler gerçekleştirilecektir. Öznitelik çıkarımının ardından model eğitme aşamasına geçilecektir.

Çalışmamızda kullanacak olduğumuz Bi-LSTM modeli, ardışık verilerde (metin, zaman serileri vb.) çalışmak için tasarlanmış bir yapay sinir ağıdır. Bu nedenle Bi-LSTM modelini, şarkı sözlerine göre şarkı türünü elde eden bir sınıflandırma modeli olarak kullanmak uygun bir seçimdir. Uygulayacağımız çözümde, Bi-LSTM modeli, Türkçe için eğitilmiş olan bir dizi fastText kelime gömme vektörünü girdi olarak alır. fastText kelime gömme vektörleri, her bir kelimenin temsil edildiği gömme vektörleridir. Daha sonra, bir dizi Bi-LSTM hücresi kullanarak ardışık verilerin işlenmesi için bir mekanizma sağlayacağız.

LSTM hücreleri, önceki durum ve girdi vektörlerinin birleşimine dayanarak, bir sonraki çıktı ve hücre durumunu üretirler. Bu işlem, bir dizi LSTM hücresi için tekrarlanır. Son olarak, Bi-LSTM hücrelerinin çıktıları, bir dizi yoğun katman kullanılarak sınıflandırma için bir tahmin yapmak için işlenir. Yoğun katmanlar, girdi vektörlerini birkaç kez işleyerek, nihai sınıflandırma sonucunu üretir.

## 4 Deneysel Değerlendirme

### 4.1 Değerlendirme Soruları

1.Müzik türleri şarkıya ait sözler vasıtasıyla ayrıştırılabilir mi? Örnek Durum: Müzik türlerini oluşturan ana unsur müziğin söyleniş biçimi ve melodiyken şarkı sözlerinin melodinin verdiği

duyguya biçimlenmesine sebep olabilir ve bu türün belirlenmesinde fayda sağlayabilir.

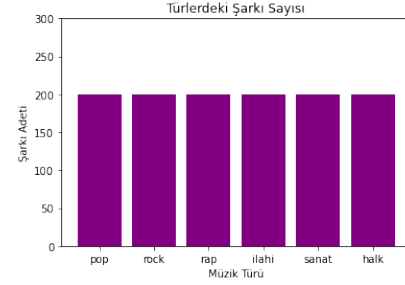


Figure 2: Şarkı Sayısı

2.Müzik türleri şarkı sözlerindeki kelimelerin uzunluklarıyla ayrıştırılabilir mi? Örnek Durum: İlahi müzik türünde kelimeler daha uzunken diğer müzik türlerinde nispeten daha yeni ve kısa kelimelerden oluşur dolayısıyla başarıyı arttırıcı bir öznitelik olabilir.

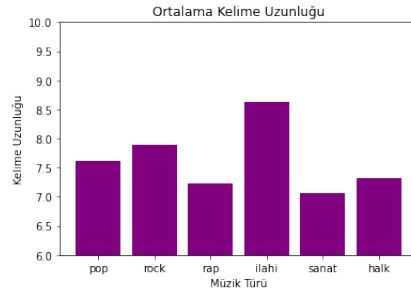


Figure 3: Ortalama Kelime Uzunlukları

3.Müzik türleri şarkı sözlerindeki kelime sayılarıyla ayrıştırılabilir mi? Örnek Durum: Rap müzik türünde şarkı sözleri uzun ve dolayısıyla birçok kelimeden oluşurken pop müzik türü daha az kelimeden oluşuyor dolayısıyla müzik türü belirlemede başarıyı arttırabilir.

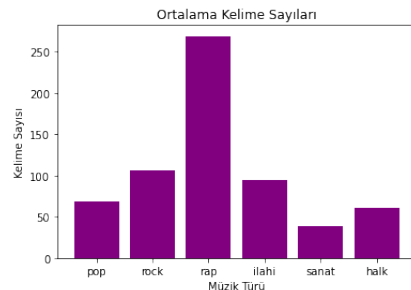


Figure 4: Ortalama Kelime Sayısı

4.Müzik türleri şarkı sözlerinde tekrar eden kelime sayıları arasında bir ilişki olabilir mi? Örnek Durum: Pop müzik türüne ait şarkı sözlerinde

<sup>9</sup><https://textblob.readthedocs.io/en/dev/>

cümle tekrarları sıklıkla yapılırken rap müzik türünde tekrara gidilmez dolayısıyla müzik türünün belirlenmesinde faydası olabilir.

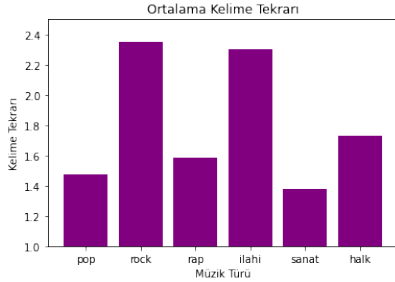


Figure 5: Ortalama Kelime Tekrarı

5. Müzik türleri şarkı sözlerinin duygu analizi sonuçlarıyla ilişkili olabilir mi? Örnek Durum: Nötr, korku, şaşkınlık, kızgınlık, üzüntü ve mutluluk duygularının şarkı sözlerinden çıkarıldık-tan sonra müzik türünü belirlemesi düşünülmüştür. Pop müzik türü daha çok üzüntü ağırlıklıyken sanat müziğinde mutluluk duygusu hakimdir.

## 4.2 Deneyisel kurulum

Bu projede, derin öğrenme modellerinin eğitimi için Keras <sup>10</sup>, NLTK <sup>11</sup> ve FastText <sup>12</sup> gibi TensorFlow <sup>13</sup> üzerine inşa edilmiş kütüphaneleri, PyTorch <sup>14</sup> gibi akademik çalışmalarda sıkça kullanılan derin öğrenme kütüphanesini ve kelime türlerini bulmak (POS-Tagging) için Türkçe Doğal Dil İşleme çalışmaları için özel olarak geliştirilmiş Zemberek <sup>15</sup> kütüphanesini ayrıca veri kümesinin işlenmesi ve görselleştirme için sklearn <sup>16</sup>, transformers <sup>17</sup>, numpy <sup>18</sup> ve pandas <sup>19</sup> gibi kütüphaneleri kullandık.

Bu projede kullanılan veri seti, çeşitli Türkçe müzik sitelerinden alınmış Türkçe şarkı sözlerinden oluşur. Veri setinde, halk, sanat, ilahi, pop, rock, rap olmak üzere 6 farklı tür için eşit sayıda örnekle, toplamda 1200 şarkı sözü bulunmaktadır. Veri seti, 0.8 oranında eğitim ve 0.2 oranında test verisi olarak ayrılmıştır. Ayrıca Bi-LSTM mod-

elinde, eğitim verisinin 0.3'ü doğrulama (validation) verisi olarak ayrılmıştır.

Projede değerlendirme ölçütleri olarak doğruluk (accuracy) ve her bir sınıf için hesaplanan F1 skorlarının ağırlıklı ortalamaları kullanılmıştır. Doğruluk ölçütü, sınıflandırma probleminde sınıfların dengeli olduğu durumlarda doğru bir etkililik ölçütü olarak kullanılabilir. Hassasiyet ölçütü, yanlış pozitiflerin önemli olduğu durumlarda tercih edilebilir. Duyarlılık ölçütü ise yanlış negatiflerin önemli olduğu durumlarda tercih edilebilir. Son olarak F1 ölçütü de sınıflandırma probleminde sınıfların dengesiz olduğu durumlarda etkililik ölçütü olarak kullanılabilir. Bu ölçütlerin kullanılmasının sebebi, farklı durumlarda farklı önceliklerin olabileceği bir sınıflandırma problemi olduğu için her bir ölçütün farklı açılardan bilgi sağlamasıdır. Örneğin, bu projedeki gibi bir müzik sınıflandırma probleminde, bazı müzik türlerinin diğerlerinden daha nadir olabileceği veya yanlış bir müzik türü tahmini yapmanın daha olumsuz sonuçlar doğurabileceği durumlar olabilir. Bu nedenle farklı ölçütlerin kullanılması, sınıflandırma probleminin performansını daha iyi anlamamıza yardımcı olmuştur.

Öncelikle Makine öğrenimi modeli olarak kullandığımız Rastgele Orman (Random Forest) modelinin parametreleri de aşağıdaki gibidir:

- maksimum derinlik (max\_depth): 30
- minimum yaprak örneği (min\_samples\_leaf): 1
- minimum örnek bölme sayısı (min\_samples\_split): 5
- tahminci sayısı (n\_estimators): 600
- rastgele durum (random\_state): 42

Projede kullandığımız derin öğrenme modeli olan Bi-LSTM modelinin katmanları ve parametre değerleri aşağıdaki gibidir:

- Bi-LSTM katmanı:
  - gizli hücre birimi (units): 128
  - dropout: 0.2
  - recurrent\_dropout: 0.2
- Yoğun katman (Dense):
  - gizli hücre birimi (units): 64
  - aktivasyon fonksiyonu (activation): 'relu'

<sup>10</sup><https://keras.io/>  
<sup>11</sup><https://www.nltk.org/>  
<sup>12</sup><https://fasttext.cc/>  
<sup>13</sup><https://www.tensorflow.org/?hl=tr>  
<sup>14</sup><https://pytorch.org/>  
<sup>15</sup><https://github.com/ahmetaa/zemberek-nlp>  
<sup>16</sup><https://scikit-learn.org/stable/>  
<sup>17</sup><https://huggingface.co/docs/transformers/index>  
<sup>18</sup><https://numpy.org/>  
<sup>19</sup><https://pandas.pydata.org/>

- Dropout katmanı:
  - rate: 0.5
- Yoğun katman (Dense):
  - çıkış sınıfı (units): 6
  - aktivasyon fonksiyonu (activation): 'soft-max'

Ayrıca, modelin eğitimi sırasında kullanılan optimize edici (optimizer) 'adam' ve kayıp fonksiyonu (loss function) 'sparse\_categorical\_crossentropy' olarak belirlenmiştir. Metrik olarak ise 'accuracy' kullanılmaktadır.

**Temel Model:** Projede temel model olarak bi-gram'lar ile SVM algoritmasını ve hassas ayar yapılmış BERT modelini kullanacağız (Devlin et al., 2019). Bu algoritmalar yapacağımız sınıflandırma işleminin sonucu açısından bize alt sınır oluşturacak ve yapacağımız geliştirmeler ile bu alt sınırı geçmeye çalışacağız.

Temel model olarak seçtiğimiz ve önceden eğitilmiş olan BERT modelini kendi verilerimize göre hassas ayar yapmak için çeşitli parametre ayarları kullandık. Bu parametreler ve değerleri aşağıda listelenmiştir:

- önceden eğitilmiş model (pre-trained model): dbmdz/bert-base-turkish-cased
- etiket sayısı (num\_labels): 6 - epoch sayısı: 4
- grup boyutu (batch size): 2
- optimize edici (optimizer): AdamW
- öğrenme oranı (learning rate): 2e-5
- epsilon: 1e-8

### 4.3 DeneySEL Sonuçlar

#### 4.3.1 Makine Öğrenmesi Modeli

Belirlediğimiz toplam Kelime Sayısı, ortalama kelime uzunlukları, ortalama kelime tekrarı, duygu analizi özneliliklerini ve şarkı sözlerinden oluşturduğumuz NGramları kullanarak Random Forest, SVM, Naive Bayes ve XGBoost makine öğrenmesi yöntemleri kullanarak sınıflandırma yapılmıştır. Sınıflandırma modelleri eğitilirken modellerin hiperparametreleri gridSearch yöntemiyle saptanmış ve optimal sonuçlar elde edilmeye çalışılmıştır. Modellerin doğruluk değerleri sırasıyla 0.77, 0.56, 0.54, 0.70 f1 skorları ise 0.75, 0.53, 0.49, 0.70 olarak karşımıza çıkmıştır. Temel Model olarak kullandığımız ve sadece NGramlar kullanılarak oluşturulan doğruluk değeri 0.35

olarak karşımıza çıkan SVM modeli yeni özniteliklerle 0.55'e çıkmış ve random forest algoritmasının 0.75'lik başarımıyla geride bırakılmıştır. Temel model olarak kullandığımız doğruluk değeri 0.85 olarak karşımıza çıkan Bert modeline göre ise en yakın makine öğrenmesi modeli 0.77 doğruluk oranıyla random forest olmuştur.

Model	Doğruluk	F1 Skoru
Random Forest	0.77	0.75
SVM	0.56	0.53
Naive Bayes	0.54	0.49
XGBoost	0.70	0.70

Table 2: Performans Ölçümleri

#### 4.3.2 Bi-LSTM Modeli

Türkçe kelimeler üzerinde eğitilmiş fastText modeli ile elde edilen kelime vektörleri LSTM modeline verildiğinde ilk başta 0.15 gibi kötü sonuçlar elde edildi. Ancak sonrasında doğrulamak için kullanılan sonuç etiketlerinin boyutu ile model parametrelerinin uygun olmadığı tespit edildi. İlk başta eğitim için kullanılan sonuç etiketleri one-hot encoding ile 6 boyutlu vektörler olarak ifade edilip LSTM modelinde kullanılan loss fonksiyonu 'categorical\_crossentropy' olarak seçildiğinde model 0.71 başarımlı gösterdi. Sonrasında sonuç etiketleri değiştirilmeden loss fonksiyonu 'sparse\_categorical\_crossentropy' olarak değiştirildiğinde 0.74 başarımlı gösterdi. Son olarak, durak kelimelerin çıkarılmadığı veri setini kullanarak modeli eğittiğimde 0.77 başarımlı elde ettim. Bu bakımdan henüz temel modellerden olan hassas ayar yapılmış BERT modelini geçemedik. Aşağıda fastText ile eğitilen Bi-LSTM modelinin tur(epoch)-doğruluk(accuracy) grafiği görülmektedir:

Model	Başarı Oranı
fastText + LSTM (one-hot encoding)	0.71
fastText + LSTM (sparse categorical crossentropy)	0.74
fastText + LSTM (durak kelimeler çıkarılmadan)	0.77

Table 3: fastText ve LSTM Model Başarımları

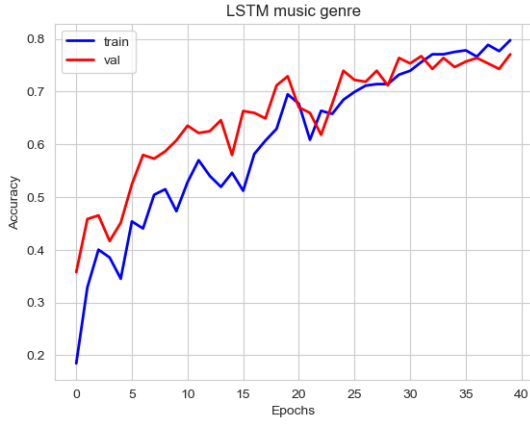


Figure 6: tur(epoch)-doğruluk(accuracy) grafiği

## 5 Sonuçlar ve Gelecek Çalışmalar

Bu Doğal Dil İşleme projesinde, farklı yöntemler kullanarak bir veri setindeki şarkı sözlerinin müzik türüne göre sınıflandırılması amaçlanmıştır. Makine öğrenmesi yöntemlerinden Random Forest, SVM, Naive Bayes ve XGBoost, derin öğrenme modeli olarak da Bi-LSTM kullanılarak modeller eğitilmiş ve en yüksek başarı skoru %77 olarak elde edilmiştir. Ayrıca, temel model olarak seçilen BERT modeli, hassas ayar yapıldığında %85 başarı elde etmiştir.

Buna karşın, çalışmamızın bazı sınırlılıkları da vardır. Örneğin, veri setimizdeki az sayıda örnek, Bi-LSTM modelinin dalgalı sonuçlar vermesine neden olmuştur. Ayrıca, eğitim ve test veri seti ayrımlarında, veri setindeki şarkıların farklı müzik türleri arasında dengeli bir şekilde dağılmadığı gözlemlenmiştir.

Gelecek çalışmalarda, özellikle veri setinin daha dengeli hale getirilmesi için daha fazla veri toplanması ve kullanılması önerilebilir. Ayrıca, öznetelik seçiminde farklı yöntemlerin denenebilmesi ve model performansının artırılması için farklı hiperparametre ayarları da denenebilir. Ayrıca, daha büyük ve karmaşık modeller kullanarak daha yüksek başarı skorları elde etmek de mümkündür.

Sonuç olarak, bu çalışma, makine öğrenmesi ve derin öğrenme yöntemlerinin müzik türü sınıflandırmasında kullanılabilirliğini göstermektedir. Gelecek çalışmalarda, daha iyi sonuçlar elde etmek için veri setinin iyileştirilmesi, farklı yöntemlerin deneyi ve daha büyük ve karmaşık modellerin kullanılması önerilebilir.

Github linki:

<https://github.com/hasanerdemak/NLP-Music-Genre-Classification>

## Referanslar

- Anna Boonyanit and Andrea Dahl. 2022. [Music genre classification using song lyrics](#).
- Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. [Bert: Pre-training of deep bidirectional transformers for language understanding](#).
- Michael Fell and Caroline Sporleder. 2014. [Lyrics-based analysis and classification of music](#). In *Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers*, pages 620–631, Dublin, Ireland. Dublin City University and Association for Computational Linguistics.
- Serhat Hızlısoy and Zekeriya Tüfekci. 2021. [Derin Öğrenme İle türkçe müziklerden müzik türü sınıflandırması](#). pages 176 – 183. Osman SAĞDIÇ.
- Florian Kleedorfer, Peter Knees, and Tim Pohle. 2008. [Oh oh oh whoah! towards automatic topic detection in song lyrics](#). pages 287–292.
- Akshi Kumar, Arjun Rajpal, and Dushyant Rathore. 2018. [Genre classification using word embeddings and deep learning](#). In *2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, pages 2142–2146.
- Cyril Laurier, Jens Grivolla, and Perfecto Herrera. 2008. [Multimodal music mood classification using audio and lyrics](#). pages 688–693.
- B. Logan, A. Kositsky, and P. Moreno. 2004. [Semantic analysis of song lyrics](#). In *2004 IEEE International Conference on Multimedia and Expo (ICME) (IEEE Cat. No.04TH8763)*, volume 2, pages 827–830 Vol.2.
- Jose Mahedero, Alvaro Martinez, Pedro Cano, Markus Koppenberger, and Fabien Gouyon. 2005. [Natural language processing of lyrics](#). pages 475–478.
- Rudolf Mayer, Robert Neumayer, and Andreas Rauber. 2008. [Combination of audio and lyrics features for genre classification in digital audio collections](#). pages 159–168.
- Rudolf Mayer and Andreas Rauber. 2011. [Music genre classification by ensembles of audio and lyrics features](#). pages 675–680.
- Cory McKay, John Burgoyne, Jason Hockman, Jordan Smith, Gabriel Vigliensoni, and Ichiro Fujinaga. 2010. [Evaluating the genre classification performance of lyrical features relative to audio, symbolic and cultural features](#). pages 213–218.
- Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg S Corrado, and Jeff Dean. 2013. [Distributed representations of words and phrases and their compositional-ity](#). In *Advances in Neural Information Processing Systems*, volume 26. Curran Associates, Inc.



- Sam Scott and Stan Matwin. 1998. Text classification using wordnet hypernyms. In *WordNet@ACL/COLING*.
- Alexandros Tsaptsinos. 2017. [Lyrics-based music genre classification using a hierarchical attention network](#).
- Dan Yang and Won-Sook Lee. 2009. [Music emotion identification from lyrics](#). In *2009 11th IEEE International Symposium on Multimedia*, pages 624–629.
- Teh Chao Ying, Shyamala C. Doraisamy, and Lili Nurliyana Abdullah. 2012. Genre and mood classification using lyric features. *2012 International Conference on Information Retrieval & Knowledge Management*, pages 260–263.
- Önder Çoban. 2017. [Turkish music genre classification using audio and lyrics features](#). volume 21, pages 322 – 331. Süleyman Demirel Üniversitesi.
- Önder Çoban and Isil Karabey Aksakalli. 2017. [Music genre classification with word and document vectors](#).
- Önder Çoban and Gülşah Tümöklü Özyer. 2016. [Music genre classification from turkish lyrics](#). In *2016 24th Signal Processing and Communication Application Conference (SIU)*, pages 101–104.

850  
851  
852  
853  
854  
855  
856  
857  
858  
859  
860  
861  
862  
863  
864  
865  
866  
867  
868  
869  
870  
871  
872  
873  
874  
875  
876  
877  
878  
879  
880  
881  
882  
883  
884  
885  
886  
887  
888  
889  
890  
891  
892  
893  
894  
895  
896  
897  
898  
899