

**BURSA TEKNİK ÜNİVERSİTESİ ❖ LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ**

**OLASILIKSAL MODELLER İLE TÜRKÇE MÜZİK TÜRLERİNİN  
SINIFLANDIRILMASI**

**YÜKSEK LİSANS TEZİ**

**Mehmet Cüneyt ÖZBALCI**

**Bilgisayar Mühendisliği Anabilim Dalı**

**Bilgisayar Mühendisliği Programı**

**OCAK 2022**

**BURSA TEKNİK ÜNİVERSİTESİ ❖ LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ**

**OLASILIKSAL MODELLER İLE TÜRKÇE MÜZİK TÜRLERİNİN  
SINIFLANDIRILMASI**

**YÜKSEK LİSANS TEZİ**

**Mehmet Cüneyt ÖZBALCI  
(19376482001)**

**Bilgisayar Mühendisliği Anabilim Dalı**

**Bilgisayar Mühendisliği Programı**

**Tez Danışmanı: Prof. Dr. Turgay Tugay BİLGİN**

**OCAK 2022**



BTÜ, Lisansüstü Eğitim Enstitüsü'nün 19376482001 numaralı Yüksek Lisans Öğrencisi Mehmet Cüneyt ÖZBALCI, ilgili yönetmeliklerin belirlediği gerekli tüm şartları yerine getirdikten sonra hazırladığı "OLASILIKSAL MODELLER İLE TÜRKÇE MÜZİK TÜRLERİNİN SINIFLANDIRILMASI" başlıklı tezini aşağıda imzaları olan jüri önünde başarı ile sunmuştur.

**Tez Danışmanı :**      **Prof. Dr. Turgay Tugay BİLGİN**      .....

Bursa Teknik Üniversitesi

**Jüri Üyeleri :**      **Doç. Dr. Gıyasettin ÖZCAN**      .....

Bursa Uludağ Üniversitesi

**Dr. Öğr. Üyesi Erdem YAVUZ**      .....

Bursa Teknik Üniversitesi

**Teslim Tarihi**      :  
**Savunma Tarihi**    : 6 Ocak 2022



20.04.2016 tarihli Resmi Gazete’de yayımlanan Lisansüstü Eğitim ve Öğretim Yönetmeliğinin 9/2 ve 22/2 maddeleri gereğince; Bu Lisansüstü teze, Bursa Teknik Üniversitesi’nin abonesi olduğu intihal yazılım programı kullanılarak Lisansüstü Eğitim Enstitüsü’nün belirlemiş olduğu ölçütlere uygun rapor alınmıştır.

## İNTİHAL BEYANI

Bu tezde görsel, işitsel ve yazılı biçimde sunulan tüm bilgi ve sonuçların akademik ve etik kurallara uyularak tarafımdan elde edildiğini, tez içinde yer alan ancak bu çalışmaya özgü olmayan tüm sonuç ve bilgileri tezde kaynak göstererek belgelediğimi, aksinin ortaya çıkması durumunda her türlü yasal sonucu kabul ettiğimi beyan ederim.

Öğrencinin Adı Soyadı: Mehmet Cüneyt ÖZBALCI

İmzası:



## ÖNSÖZ

Bu tez çalışmamda; emeklerini hiçbir zaman esirgemeyen, her konuda bana destek ve motive olan tez danışmanım Prof. Dr. Turgay Tugay BİLGİN hocama sonsuz teşekkürlerimi sunuyorum.

Ocak 2022

Mehmet Cüneyt ÖZBALCI  
(Araştırma Görevlisi)



## İÇİNDEKİLER

### Sayfa

ÖNSÖZ .....	vi
İÇİNDEKİLER .....	vii
KISALTMALAR .....	ix
SEMBOLLER .....	xi
ÇİZELGE LİSTESİ.....	xii
ŞEKİL LİSTESİ.....	xiv
ÖZET.....	xvi
SUMMARY .....	xvii
<b>1. GİRİŞ .....</b>	<b>1</b>
1.1 Tezin Amacı .....	3
1.2 Tezin Katkısı .....	3
1.3 Literatür Araştırması .....	4
<b>2. MÜZİKTE SES SİNYALLERİNİN İŞLENMESİ.....</b>	<b>8</b>
2.1 Örneklem Oranı (Sample Rate).....	8
2.2 Frekans ve Genlik.....	9
<b>3. MAKİNE ÖĞRENMESİ ALGORİTMALARI VE SINIFLANDIRMA.....</b>	<b>10</b>
3.1 Makine Öğrenmesi .....	10
3.1.1 Denetimli öğrenme (Supervised learning) .....	11
3.1.2 Denetimsiz öğrenme (Unsupervised learning).....	11
3.1.3 Yarı denetimli öğrenme (Semi supervised learning) .....	11
3.1.4 Aktif öğrenme (Active learning).....	12
3.1.5 Aşırı öğrenme (Overfitting) .....	12
3.1.6 Sınıflandırma.....	12
3.1.7 Karmaşıklık matrisleri.....	13
3.1.8 Sınıflandırma başarımları ile ilgili metrikler .....	14
3.2 K-NN Algoritması .....	15
3.2.1 K-NN'de mesafe hesaplama yöntemleri.....	17
3.3 SVM Algoritması .....	18
3.3.1 Doğrusal SVM .....	18
3.3.2 Doğrusal olmayan SVM.....	19
3.3.3 Çekirdek fonksiyonları.....	20
3.3.3.1 Polinomal çekirdek fonksiyonu.....	20
3.3.3.2 Radyal tabanlı çekirdek fonksiyonu.....	21
3.3.3.3 Lineer çekirdek fonksiyonu .....	22
3.4 Lojistik Regresyon .....	22
<b>4. ÖZİNİTELİK ÇIKARIMI.....</b>	<b>25</b>
4.1 Mel Frekans Kepstral Katsayıları (MFCCs) .....	25
4.1.1 Ön vurgulama.....	26
4.1.2 Çerçeveleme.....	26
4.1.3 Pencereleme .....	27
4.1.4 Hızlı Fourier dönüşümü (FFT).....	28

4.1.5 Mel filtre bankası .....	29
4.1.6 Ayrık kosinüs dönüşümü (DCT).....	30
4.2 Sıfır Geçiş Oranı (Zero Crossing Rate).....	31
4.3 İzgesel Merkez (Spectral Centroid).....	32
4.4 İzgesel Bant Geniřlięi (Spectral Bandwidth).....	33
4.5 İzgesel Azalma (Spectral Rolloff).....	34
4.6 Chroma STFT.....	35
4.7 Ortalama Karesel Hata (RMSE).....	35
<b>5. BULGULAR VE YORUMLAR .....</b>	<b>37</b>
5.1 Veri Seti.....	37
5.2 Wav Müzik Formatı .....	37
5.3 Tezin Akıř Őeması.....	38
5.4 K-NN Algoritması ile Müzik Türlerinin Sınıflandırılması .....	39
5.5 SVM Algoritması ile Müzik Türlerinin Sınıflandırılması.....	50
5.6 Lojistik Regresyon Algoritması ile Müzik Türlerinin Sınıflandırılması.....	61
<b>6. SONUÇLAR VE ÖNERİLER .....</b>	<b>65</b>
<b>KAYNAKLAR .....</b>	<b>67</b>
<b>EKLER.....</b>	<b>73</b>
<b>ÖZGEÇMİŐ.....</b>	<b>74</b>

## KISALTMALAR

<b>AUC-ROC</b>	: Eğri Altındaki Alan - Alıcı Çalışma Karakteristikleri
<b>BiLSTM</b>	: Çift Yönlü Uzun Kısa Süreli Bellek
<b>CNN</b>	: Evrimsel Sinir Ağı
<b>CRNN</b>	: RNN ve CNN'i Kapsayan 2 Katmanlı Mimari
<b>csv</b>	: Virgül ile Ayrılmış Değerler
<b>dB SPL</b>	: Ses Basıncı Seviyesi
<b>DCT</b>	: Ayrık Kosinüs Dönüşümü
<b>DFT</b>	: Ayrık Fourier Dönüşümü
<b>DNN</b>	: Derin Anlama Ağı
<b>FFT</b>	: Hızlı Fourier Dönüşümü
<b>FN</b>	: Yanlış Negatif
<b>FNN</b>	: İleri Beslemeli Sinir Ağı
<b>FP</b>	: Yanlış Pozitif
<b>GRU</b>	: Tekrarlayan Kapı Birimi
<b>IBM</b>	: Uluslararası İş Makineleri
<b>kHz</b>	: Kiloherz
<b>K-NN</b>	: K-En Yakın Komşular
<b>LMD</b>	: Latin Müzik Veritabanı
<b>LSTM</b>	: Uzun-Kısa Vadeli Bellek
<b>MDS</b>	: Çok Boyutlu Ölçekleme
<b>MFCCs</b>	: Mel Frekans Kepstral Katsayıları
<b>MIDI</b>	: Müzik Enstürmanları Dijital Arabirimi
<b>MIR</b>	: Müzik Bilgi Erişimi
<b>mLSTM</b>	: Çarpımsal Uzun Kısa Süreli Bellek
<b>OSH</b>	: Optimum Ayrılabilir Hiper Düzlem
<b>PCA</b>	: Temel Bileşen Analizi
<b>PWED</b>	: Algısal Ağırlıklı Öklid Mesafesi
<b>RMSE</b>	: Ortalama Karesel Hata
<b>RNN</b>	: Yinelemeli Sinir Ağı
<b>STFT</b>	: Kısa Zamanlı Fourier Dönüşümü
<b>SVM</b>	: Destek Vektör Makineleri

<b>TN</b>	: Gerçek Negatif
<b>TP</b>	: Gerçek Pozitif
<b>tSNE</b>	: t-Dağıtılmış Stokastik Komşu Gömme
<b>wav</b>	: Dalga Formu Ses Dosyası Formatı
<b>zcr</b>	: Sıfır Geçiş Oranı



## SEMBOLLER

<b>a</b>	: Filtre katsayısı
<b>c</b>	: C sabiti
<b>d</b>	: Polinom derecesi
<b>e</b>	: Euler sayısı
<b>p</b>	: İzgesel bant genişliğinin derecesi
<b>t</b>	: Zaman
<b>T, <math>1_{R&lt;0}</math></b>	: İşaretçi fonksiyon
<b><math>\sigma</math></b>	: Sigma parametresi

## ÇİZELGE LİSTESİ

### Sayfa

<b>Çizelge 5.1</b> : K-NN ile öklid mesafe fonksiyonu kullanılarak elde edilen en yüksek başarıımı veren öznitelik seçimi ile sınıflandırılan türlerin başarımlarını gösteren metrikler. ....	<b>46</b>
<b>Çizelge 5.2</b> : K-NN ile manhattan mesafe fonksiyonu kullanılarak elde edilen en yüksek başarıımı veren öznitelik seçimi ile sınıflandırılan türlerin başarımlarını gösteren metrikler .....	<b>46</b>
<b>Çizelge 5.3</b> : K-NN ile chebyshev mesafe fonksiyonu kullanılarak elde edilen en yüksek başarıımı veren öznitelik seçimi ile sınıflandırılan türlerin başarımlarını gösteren metrikler .....	<b>47</b>
<b>Çizelge 5.4</b> : K-NN ile arabesk müzik türü için mesafe fonksiyonlarına ait metriklerin en yüksek başarımları.....	<b>47</b>
<b>Çizelge 5.5</b> : K-NN ile dini müzik türü için mesafe fonksiyonlarına ait metriklerin en yüksek başarımları .....	<b>48</b>
<b>Çizelge 5.6</b> : K-NN ile klasik müzik türü için mesafe fonksiyonlarına ait metriklerin en yüksek başarımları .....	<b>48</b>
<b>Çizelge 5.7</b> : K-NN ile pop müzik türü için mesafe fonksiyonlarına ait metriklerin en yüksek başarımları .....	<b>48</b>
<b>Çizelge 5.8</b> : K-NN ile rap müzik türü için mesafe fonksiyonlarına ait metriklerin en yüksek başarımları .....	<b>49</b>
<b>Çizelge 5.9</b> : K-NN ile rock müzik türü için mesafe fonksiyonlarına ait metriklerin en yüksek başarımları .....	<b>49</b>
<b>Çizelge 5.10</b> : SVM ile radyal tabanlı çekirdek fonksiyonu kullanılarak elde edilen en yüksek başarıımı veren öznitelik seçimi ile sınıflandırılan türlerin başarımlarını gösteren metrikler .....	<b>57</b>
<b>Çizelge 5.11</b> : SVM ile polinomal çekirdek fonksiyonu kullanılarak elde edilen en yüksek başarıımı veren öznitelik seçimi ile sınıflandırılan türlerin başarımlarını gösteren metrikler .....	<b>57</b>
<b>Çizelge 5.12</b> : SVM ile lineer çekirdek fonksiyonu kullanılarak elde edilen en yüksek başarıımı veren öznitelik seçimi ile sınıflandırılan türlerin başarımlarını gösteren metrikler .....	<b>58</b>
<b>Çizelge 5.13</b> : SVM ile arabesk müzik türü için çekirdek fonksiyonlarının en yüksek başarımları .....	<b>58</b>
<b>Çizelge 5.14</b> : SVM ile dini müzik türü için çekirdek fonksiyonlarının en yüksek başarımları .....	<b>59</b>
<b>Çizelge 5.15</b> : SVM ile klasik müzik türü için çekirdek fonksiyonlarının en yüksek başarımları .....	<b>59</b>
<b>Çizelge 5.16</b> : SVM ile pop müzik türü için çekirdek fonksiyonlarının en yüksek başarımları .....	<b>60</b>
<b>Çizelge 5.17</b> : SVM ile rap müzik türü için çekirdek fonksiyonlarının en yüksek başarımları .....	<b>60</b>
<b>Çizelge 5.18</b> : SVM ile rock müzik türü için çekirdek fonksiyonlarının en yüksek başarımları .....	<b>60</b>

<b>Çizelge 5.19</b> : Lojistik regresyon algoritması ile elde edilen en yüksek başarıyı veren öz nitelik seçimi ile sınıflandırılan türlerin başarımlarını gösteren metrikler.....	<b>64</b>
--	-----------



## ŞEKİL LİSTESİ

### Sayfa

Şekil 2.1 : Veri setine ait bir parçanın python ile elde edilmiş genlik-zaman grafiği .	8
Şekil 2.2 : Çeşitli ses sinyallerine ait frekans-genlik grafikleri .....	9
Şekil 3.1 : Karmaşıklık matrisi .....	14
Şekil 3.2 : K değeri 3 ve 7 seçildiğinde K-NN'nin en yakın komşulara yaklaşımı...	16
Şekil 3.3 : Doğrusal SVM'e ait çeşitli hiper düzlemler .....	19
Şekil 3.4 : Doğrusal SVM'e ait optimum hiper düzlem .....	19
Şekil 3.5 : Doğrusal olmayan SVM'e ait optimal bir hiper düzlem .....	20
Şekil 3.6 : 2. derece polinomal çekirdek fonksiyonunu ile verilerin ayrıştırılması ...	21
Şekil 3.7 : Lojistik regresyon ve lineer olasılık modellerine ait eğriler.....	23
Şekil 3.8 : Sigmoid fonksiyonuna ait eğri .....	23
Şekil 4.1 : MFCCs Öznitelik Çıkarımı Blok Diyagramı .....	25
Şekil 4.2 : Veri setine ait bir parçadan ön vurgulama ile elde edilen spektrogram ...	26
Şekil 4.3 : Çerçevelerin örtüştürülmesi .....	27
Şekil 4.4 : 400 örnek için çeşitli pencereleme fonksiyonlarına ait eğriler.....	28
Şekil 4.5 : FFT'ye ait zaman-frekans grafiği.....	29
Şekil 4.6 : 20 Mel filtresi kullanılarak elde edilen bir Mel filtre bankası örneği .....	30
Şekil 4.7 : Ayrık kosinüs dönüşümü ile elde edilen görsel.....	31
Şekil 4.8 : Ses sinyalinden çıkarılmış sıfır geçiş oranı grafiği örneği .....	32
Şekil 4.9 : Sese ait sinyalden izgesel merkezin elde edilmesi .....	33
Şekil 4.10 : İzgesel bant genişliği grafiksel gösterimi .....	34
Şekil 4.11 : $q(t)$ fonksiyonu için izgesel azalmanın grafiği .....	35
Şekil 5.1 : Çeşitli ses sinyallerine ait frekans-genlik grafikleri .....	38
Şekil 5.2 : K-NN algoritması ile mesafe fonksiyonu öklid ve K değeri 10 olduğu durumda elde edilen doğruluk oranları .....	40
Şekil 5.3 : K-NN algoritması ile $K=10$ alınarak ve öklid mesafe fonksiyonu kullanılarak elde edilen karmaşıklık matrisleri .....	41
Şekil 5.4 : K-NN algoritması ile mesafe fonksiyonu “manhattan” ve K değeri 10 olduğu durumda elde edilen doğruluk oranları .....	42
Şekil 5.5 : K-NN algoritması ile $K=10$ alınarak ve “manhattan” mesafe fonksiyonu kullanılarak elde edilen karmaşıklık matrisleri .....	43
Şekil 5.6 : K-NN algoritması ile mesafe fonksiyonu “chebyshev” ve K değeri 7 olduğu durumda elde edilen doğruluk oranları .....	44
Şekil 5.7 : K-NN algoritması ile $K=7$ alınarak ve “chebyshev” mesafe fonksiyonu kullanılarak elde edilen karmaşıklık matrisleri .....	45
Şekil 5.8 : K-NN algoritması ile seçilen mesafe fonksiyonuna göre elde edilen en yüksek doğruluk oranları.....	50
Şekil 5.9 : SVM algoritması ile radyal tabanlı çekirdek fonksiyonuna ait doğruluk oranları .....	51
Şekil 5.10 : SVM algoritması ile radyal tabanlı çekirdek fonksiyonu kullanılarak elde edilen karmaşıklık matrisleri.....	52
Şekil 5.11 : SVM algoritması ile polinomal çekirdek fonksiyonuna ait doğruluk oranları .....	53

<b>Şekil 5.12</b> : SVM algoritması ile polinomal çekirdek fonksiyonu kullanılarak elde edilen karmaşıklık matrisleri.....	<b>54</b>
<b>Şekil 5.13</b> : SVM algoritması ile lineer çekirdek fonksiyonuna ait doğruluk oranları .....	<b>55</b>
<b>Şekil 5.14</b> : SVM algoritması ile lineer çekirdek fonksiyonu kullanılarak elde edilen karmaşıklık matrisleri.....	<b>56</b>
<b>Şekil 5.15</b> : SVM algoritması ile uygulanan çekirdek fonksiyonları için elde edilen en yüksek doğruluk oranları.....	<b>61</b>
<b>Şekil 5.16</b> : Lojistik regresyon algoritması ile elde edilen doğruluk oranları .....	<b>62</b>
<b>Şekil 5.17</b> : Lojistik regresyon algoritması ile elde edilen karmaşıklık matrisleri ....	<b>63</b>



## OLASILIKSAL MODELLER İLE TÜRKÇE MÜZİK TÜRLERİNİN SINIFLANDIRILMASI

### ÖZET

Müzikle ilgili çalışmalar günümüzde oldukça yoğun bir şekilde ilgi görmektedir. Müzik türü sınıflandırması, müzik bilgisi erişimindeki (MIR) temel problemlerden birisidir. Müzik türü sınıflandırması; makine öğrenmesi ve derin öğrenme algoritmalarının giderek geliştirilmesi ve yaygın bir şekilde kullanılması ile oldukça popüler bir çalışma alanı haline gelmiştir. Makine öğrenmesi algoritmaları ile belirli bir seviyeye kadar geliştirilen müzikten ses analizi derin öğrenme algoritmalarının kullanılmaya başlaması ile birlikte çok daha ileri seviyelere taşınmış ve oldukça başarılı sonuçlar verebilen üretici modeller geliştirilmeye başlanmıştır. Müzik analizinde ve müzik sentezinde bu algoritmalar oldukça sık kullanılmaktadır. Ses işleme, görüntü işleme, sinyal işleme gibi yöntemler ile öğrenme algoritmaları kullanılarak çeşitli yaklaşımlarla müzik sınıflandırması gerçekleştirilmektedir. Müzik parçalarına ait veriler çeşitli araçlarla elde edilerek bu verilerden anlamlı bilgiler elde edilip müzik türlerinin sınıflandırılmasında kullanılmaktadır. Müzik parçalarından veya ses dosyalarından çeşitli öznitelikler elde etmek için oldukça gelişmiş araçlar mevcuttur. Bunun yanı sıra programlama dillerine ait kütüphanelerle ihtiyaç duyulan verilerin türlerine göre doğru özellikler belirlenerek ve doğru sınıflandırıcıyı, doğru parametrelerle modelleyerek müzik türü sınıflandırmasında başarılı sonuçlar elde edilebilmektedir. Çoğu durumda, verilerden ne kadar fazla özellik çıkarılırsa o kadar başarılı sonuçlar alınması muhtemel olacaktır. Doğru özellik seçimi de sınıflandırma başarımı için çok önemli bir ölçüttür.

Gerek akademik alanda gerekse ticari alanda ilgili yöntemler kullanılarak çeşitli çalışmalar yürütülmektedir. Müzik bilgi erişiminde, müzik sınıflandırmasının yeniliklere açık bir çalışma alanı olduğu ve yakın gelecekte ilgili çalışmaların giderek daha da artacağı düşünülmektedir.

Bu tez çalışmasında SVM, K-NN ve Lojistik Regresyon algoritmaları ile ses sinyallerine ait özellikler kullanılarak Türkçe müziklerin sınıflandırması çalışması yapılmıştır. Sınıflandırmada kullanılan algoritmaların müzik türü sınıflandırmasındaki çeşitli öznitelikler altındaki başarımlarının karşılaştırmalı analizi yapılmıştır. Her bir özneliğin başarıma olan etkilerinin karşılaştırmalı analizleri sunulmuştur. Veri setine ait; arabesk, pop, rock, dini, rap ve klasik müzik olmak üzere 6 farklı türe müzik sınıflandırması yapılmıştır. Veri seti, her bir türe ait 100'er parça olmak üzere tamamı Türkçe müzik parçalarından derlenen toplamda 600 adet wav uzantılı parçadan oluşmaktadır. Elde edilen sonuçlarda 3 farklı sınıflandırma algoritmasından SVM algoritması en başarılı sonuçları vermiştir. Elde edilen en yüksek doğruluk oranı %78.65 olarak görülmüştür.

**Anahtar kelimeler:** Müzik Madenciliği, Makine Öğrenmesi, Sinyal İşleme

# CLASSIFICATION OF TURKISH MUSIC GENRES WITH PROBABILISTIC MODELS

## SUMMARY

Studies on music attract a lot of attention today. Music genre classification is one of the main problems in music information retrieval (MIR). Music genre classification; it has become a very popular field of study with the progressive development and widespread use of machine learning and deep learning algorithms. Music-to-sound analysis, which was developed to a certain level with machine learning algorithms, has been moved to a much more advanced level with the use of deep learning algorithms, and generative models that can give very successful results have begun to be developed. These algorithms are used quite frequently in music analysis and music synthesis. Music classification is carried out with various approaches using methods such as sound processing, image processing, signal processing and learning algorithms. The data of the music pieces are obtained by various tools and meaningful information is obtained from these data and used in the classification of music genres. Highly advanced tools are available to extract various attributes from music tracks or audio files. In addition, successful results in music genre classification can be obtained by determining the right features according to the types of data we need with the libraries of programming languages and modeling the right classifier with the right parameters. In most cases, the more features we extract from the data, the more successful we will be. However, correct feature selection is also a very important detail for classification performance.

Various studies are carried out using relevant methods both in the academic field and in the commercial field. In music information access, it is thought that music classification is an open field of study and related studies will increase in the near future.

In this thesis study, classification of Turkish music was carried out by using SVM, K-NN and Logistic Regression algorithms and features of audio signals.

A comparative analysis of the performances of the algorithms used in classification under various attributes in music genre classification has been made. Comparative analyzes of the effects of each attribute on performance are presented. Belonging to the data set; music classification was made into 6 different genres as arabesque, pop, rock, religious, rap and classical music. The data set consists of 600 pieces of wav extension, file all compiled from Turkish music pieces, 100 pieces of each genre. In the results obtained, SVM algorithm gave the most successful results among 3 different classification algorithms. The highest accuracy rate obtained as 78.65%.

**Keywords:** Music Mining, Machine Learning, Signal Processing

## 1. GİRİŞ

Müzik; birbiriyle uyumlu notaların bir araya gelmesiyle oluşan içerisinde ahenk barındıran anlamlı sesler bütünüdür. Ahenk barındıran tüm sesler; deniz sesi, su şırlıtısı, kuş sesi, yağmur sesi ve hatta rüzgâr sesi aslında birer müzik parçası olmakla beraber müzik parçalarında eşlik olarak da kullanılabilir. Müzik; insanın duygularını, düşüncelerini ve hislerini ifade etmesi için etkili bir araçtır. Geçmişten günümüze birçok farklı kültürden ve birçok farklı toplumdaki çeşitli müzikal eserlere ulaşmak mümkündür. Her toplumun benimsediği kendine ait bir müzik kültürü vardır. Müzik kültürlerinin çeşitliliği ve genişliği müziği kategorize etme gereğini doğurmuştur. Kişilerin bireysel algıları ile müziği kategorize edebilmesi mümkündür. Bu kategorizasyon için gelişen teknoloji ile birlikte makineler de anlamlı sonuçlar üretebilmeye başlamıştır.

Çok geniş veri setlerinde etiketli veya etiketsiz müzikler için bir kategorizasyon için makine öğrenmesi veya derin öğrenme algoritmalarını kullanmak etkili sonuçlar vermektedir. Veriler sınıflandırılırken genellikle ham veriden anlamlı özellikler çıkarılıp bu özellikler kullanılarak hangi ölçüte göre etiketleme işlemi yapılacaksa ona göre bir etiketleme işlemi yapılmaktadır.

İnsanoğlunun müzikle olan ilişkisi dönemden döneme, kültürden kültüre farklılık gösterse de tarihin her döneminde var olmuştur. Toplumların benimsediği, kültürünü yansıtan bir müzik karakteristiği vardır. İlkel toplumlardan günümüze kadar çeşitli enstrümanlar kullanılarak müzik kültürü kuşaktan kuşağa aktarılmıştır. Müziğin bu denli vazgeçilmez olması günümüzde ilgili çalışmaların da önünü açmıştır. Müzikle ilgili çalışmalar son zamanlarda oldukça yoğun bir şekilde ilgi görmektedir. Müziğin sınıflandırılması da bu çalışmalardan birisidir. Müzik türü sınıflandırması, müzik bilgisi erişiminde (MIR) temel problemlerden bir tanesidir. Müzik sınıflandırması başlıca sanatçıya göre, şarkı sözlerine göre, melodiye göre, tempoya göre, tınıya göre, enstrümana göre, kişide uyandırdığı duyguya göre müzik sınıflandırması olarak ele alınmaktadır.

Müziği; iki temel başlıkta toplamak mümkündür. Bunlar saf müzik ve vokal müziktir. Saf müzik; sadece enstrümantal müzik içeren müzik olarak tanımlanırken, vokal müzik; hem vokal hem de ses içeren müzik olarak tanımlanmaktadır [1]. Müzik türü sınıflandırmasında kullanılacak olan müziğin türü elde edilecek başarıma doğrudan etki edeceğinden ötürü önem arz etmektedir.

Müziğin var olması için ses sinyallerinin üretilmesi gerekmektedir. Bu sinyaller insan sesine ait veya enstrümanlara ait sinyallerdir. Her bir insanın sesinin farklı bir oktav aralığı ve her bir enstrümanın kendisine ait bir tınısı vardır. Hatta aynı enstrümanın farklı türevleri veya modelleri arasında dahi ciddi tını farklılıkları olabilmektedir. İşte bütün bu çeşitlilikte ses sinyallerinden özellik çıkarmak müzik bilgi erişiminde oldukça değerlidir. Ses sinyallerinden çeşitli özellikler çıkarmak mümkündür. Ses sinyalinden elde edilen özellikler ile müzik sınıflandırması yaygın olarak kullanılan bir yöntemdir.

Müzik sınıflandırması özellikle dijital müzik platformları için ayrı bir önem arz etmektedir. Çeşitli dijital müzik platformları kullanıcılarına yönelik hizmet kalitesini artırmak için daha önce kullanıcıların dinlediği şarkıları baz alarak benzer türde öneriler sunabilmek adına bazı öneri yöntemleri geliştirmiştir. Bu noktada derin öğrenme algoritmaları ve makine öğrenmesi teknikleri kritik bir rol oynamaktadır. Akademik alanda ise son yıllarda artan bir ivme ile müzik sınıflandırması ile ilgili çalışmalar yapılmaktadır. Mevcut verilerin giderek artması ve sistematik bir biçimde depolanabilmesi sayesinde ilgili çalışmaların önü açılmıştır [2]. Çeşitli türden müzik parçalarından oluşan birçok veri seti müzik ile ilgili çalışmalarda yaygın olarak kullanılmaktadır. Müzik sınıflandırma probleminde olasılıksal modellerin kullanımı oldukça yaygındır. Destek vektör makineleri (support vector machine - SVM), karar ağaçları (decision trees), K-en yakın komşular (K-nearest neighbors - K-NN), naive bayes, rassal orman (random forest), lojistik regresyon (logistic regression) gibi algoritmaların yanı sıra, evrimsel sinir ağları (convolutional neural networks - CNN), yinelenen sinir ağları (recurrent neural networks - RNN) ve uzun-kısa vadeli bellek (long-short term memory - LSTM) gibi derin ağların da günümüzde artan hesaplama kabiliyetleri sonrasında popüler hale geldiği ve giderek yaygın bir şekilde kullanıldığı görülmektedir. Müzik türü sınıflandırmasında seçilecek sınıflandırma yöntemine göre birçok parametreyi aynı anda göz önünde bulundurmanız gerekebilmektedir. Bunlar; enstrüman, tempo, tını, ritim, ezgi, melodi, perde, sanatçı, süre, frekans gibi

parametreler olabilmektedir. Müzik bilgi erişiminde, müzik sınıflandırması birçok farklı disiplinle iç içedir. Bunlardan başlıca öne çıkanlar; doğal dil işleme, sinyal işleme, makine öğrenmesi, derin öğrenme ve veri madenciliğidir.

### **1.1 Tezin Amacı**

Bu tez çalışmasının amacı; Türkçe müzikler üzerinde K-NN ve SVM ve lojistik regresyon algoritmalarının belirlenen öznitelikler ile sınıflandırma başarımlarının gözlemlenerek karşılaştırmalı analizini gerçekleştirmektir. İlgili algoritmalarla ve belirlenen özniteliklerin ayrı ayrı kullanılarak başarımlarının tespit edilip en yüksek başarıma hangi öznitelikler kullanılarak ulaşılabileceğinin gösterilmesi ve hangi özniteliklerin başarıma ne kadar etki ettiğinin gözlemlenmesi hedeflenmiştir. Türkçe müziğin tercih edilmesindeki amaç literatürdeki Türkçe müzik üzerine yapılan çalışmaların oldukça kısıtlı olması ve Batı müziğine göre farklı bir yapısı ve kendine özgü karakteristiği olan Türk müziğinin sınıflandırma başarımları gözlemlenip kullanılan Türkçe veri seti üzerinden hem türler arası sonuçların hem de algoritmalar arasındaki sonuçların karşılaştırılmasıdır. Bu amaçla tamamı Türkçe müziklerden oluşan bir veri seti hazırlanıp bu veri seti kullanılarak müzik türü sınıflandırması gerçekleştirilmesi hedeflenmiştir.

### **1.2 Tezin Katkısı**

Yapılan çalışmada farklı makine öğrenmesi algoritmalarının belirlenen öznitelikler altında başarımlarının karşılaştırılarak Türk müziğinde belirlenen türler kapsamında hangi algoritmanın hangi öznitelik ile birlikte ne şekilde performans gösterdiği gözlemlenerek karşılaştırmalı analizleri sunulmuştur. Tüm özniteliklerin tek tek müzik türü sınıflandırmasında başarıma olan etkileri gösterilmiştir. Bu çalışmanın önemli katkılarından bir tanesi de hazırlanan veri setidir. Tamamı Türkçe müzik parçalarından oluşan veri seti, tüm parçalar tek tek ayıklanıp işlenmeye hazır hale getirilerek derlenmiştir.

### 1.3 Literatür Araştırması

Müzik sınıflandırma konusunda gerçekleştirilen literatür araştırması sonucunda birçok araştırmacı tarafından gerçekleştirilmiş çalışmalara ulaşılmıştır. Bu bölümde bu çalışmalar özetlenmiştir.

Vishnupriya ve Meenakshi'nin çalışmasında CNN ile müzik türü sınıflandırılmıştır. Öznitelik çıkarımında Mel frekans kepstral katsayıları (MFCCs) kullanılmıştır. GTZAN veri seti kullanılarak 10 alt kategoriye müzik sınıflandırılması yapılmıştır. MFCCs özelliği kullanılarak %76 gibi bir doğruluk oranı elde edilmiştir [3].

Durdağ ve Erdoğan'ın çalışmasında müziklerin renklerinin belirlenip bu sayede müzik parçalarının sınıflandırılması hedeflemiştir. Müzik türlerinden seçilen parçalar renkli görüntülere dönüştürülmüştür. Renkli görüntüler bazı ön işlemlerle elde edilmiştir. Sınıflandırma algoritması olarak CNN, veri seti olarak ise GTZAN kullanılmıştır. Türkçe müzikler ile eğitilen ağda, yaklaşık %60 gibi bir doğruluk elde edilmiştir [4].

Aguiar ve arkadaşlarının çalışmasında LMD (Latin müzik veritabanı) veri setinden alınan 3227 parça ile 10 farklı kategoride otomatik müzik türü sınıflandırması yapılmıştır. Sınıflandırmada CNN algoritması kullanılmıştır. İlk olarak dijital ses sinyali spektrogramlara dönüştürülüp ardından elde edilen görüntüler parçalara ayrılıp yeniden boyutlandırılmıştır. Veriyi işleme stratejilerinden; gürültü ekleme, perde kaydırma, ses yüksekliği değişimi ve zaman uzatma seçilip uygulanmıştır. Elde edilen en iyi doğruluk, yalnızca CNN'ye dayalı yaklaşımla ve tek tonlu perde kaydırma ile %89.45 olarak tespit edilmiştir [5].

Zhao ve arkadaşlarının çalışmasında RNN, FNN ve geliştirilmiş FNN algoritmalarının sınıflandırmadaki karşılaştırmalı analizi yapılarak MIDI tipinde müzik dosyalarının 5 türde duyguya göre sınıflandırması yapılmıştır. Duygu tespitinde 15 lisans öğrencisi seçilerek deneyler yapılmıştır. Veri seti Youtube'dan elde edilmiştir. Deneysel sonuçlarda en yüksek başarımla geliştirilmiş RNN ile %75.4 olarak gözlemlenmiştir [6].

Panwar ve arkadaşlarının çalışmasında derin öznitelik öğrenme yöntemleri kullanılarak CRNN (RNN ve CNN'i kapsayan 2 katmanlı mimari) ile müzik sınıflandırması yapılmıştır. 50 adet benzer alt etiket birleştirilerek bir sınıflandırma mekanizması oluşturulmuştur. Veriyi ön işlemede log amplitude Mel-spectrogram,

STFT ve FFT yöntemleri uygulanmıştır. Deneysel sonuçlarda AUC-ROC indeksi 0.893 olarak hesaplanmıştır [7].

Guo ve arkadaşlarının çalışmasında çift yönlü RNN yapısı kullanılarak müzik melodisi sınıflandırılmıştır. Veri setini oluşturmak için Wikifonia'daki orijinal MusicXML örnekleri (yalnızca melodili), 16 bar MIDI dosyalarına ayrıştırılmış olup sınıflandırma için 3 farklı model kullanılmıştır. Bunlar: çift yönlü uzun kısa süreli bellek (BiLSTM) modeli, çarpımsal uzun kısa süreli bellek (mLSTM) modeli ve tekrarlayan kapı birimi (GRU) modelidir. Modeller arasında performans açısından büyük farklar gözlemlenmemiştir [8].

Jawaharlal Nehru ve Jothilakshmi'nin çalışmasında müzik türü sınıflandırması için derin anlama ağı (DNN) algoritmasını kullanarak dört farklı kategoride bir sınıflandırma yapılmıştır. Ham ses sinyalinden çeşitli aşamalar sonucunda MFCCs özellikleri elde edilmiştir. 400 ses dosyasının her birinin 60 saniyelik kısımları baz alınarak DNN algoritması ile elde edilen sınıflandırmada doğruluk oranı %97.8 olarak ölçülmüştür [9].

Kızrak ve Bolat'ın çalışmasında 6 Klasik Türk Müziği makamının derin anlama ağ yapısı ile sınıflandırılması üzerine çalışılmıştır. Öznitelik olarak Mel frekans kepstral katsayıları ve delta Mel frekans kepstral katsayıları, sınıflayıcı olarak derin anlama ağları kullanılmıştır. En yüksek başarımlı Mel frekans kepstral katsayıları ve derin anlama ağı ile %92.70 olarak elde edilmiştir [10].

Ritter ve arkadaşlarının çalışmasında kullanıcıların müzik tercihlerinin otomatik olarak yapılandırmasına olanak tanıyan kişiselleştirilmiş etkileşimli bir görsel müzik sınıflandırma aracı sunulmuştur. Sınıflandırmada kullanılacak algoritmayı kullanıcı seçmektedir. Bu algoritmalar; random forest, C4.5, k-star, naive bayes, simple logistic regression (basit lojistik regresyon), multilayer perceptron (çok katmanlı algılayıcı), SVM ve AdaBoost olabilmektedir. 3D görüntüyü 2D görüntüye indirgeyen PCA, tSNE ve MDS teknikleri ile parçaların renklere ayrıştırılıp (mavi, mor, yeşil ve sarı) 2D görünümünün elde edilip böylece hangi sınıfa ait olduğu bilgisi elde edilmiş ve sınıflandırma işlemi gerçekleştirilmiştir [11].

Silla ve arkadaşlarının çalışmasında otomatik müzik türü sınıflandırması problemi için bir topluluk yaklaşımı sunulmuştur. Sınıflandırmada; naive bayes, decision tree, SVM ve çok katmanlı algılayıcı kullanılmıştır. Müzik bölümleri zaman dilimlerine göre

başından ortasından ve sonundan ayrıştırılmıştır. Nihâî müziği üretmek için elde edilen sonuçlar birleştirilmiştir. Nihâî sınıflandırma, bir kombinasyon prosedürüne göre bireysel sonuçlar kümesinden elde edilmiştir. Kullanılan veri seti LMD'dir. Çoğu durumda global ve bireysel kesim sınıflandırıcılardan elde edilenlere göre önerilen topluluk yaklaşımı daha iyi sonuçlar vermektedir [12].

Karatana ve Yıldız'ın çalışmasında veri setine ilk olarak sinyal işleme uygulanıp her bir parça için öznitelik çıkarma işlemi uygulanarak SVM, K-NN, rassal orman ve yapay sinir ağları algoritmaları kullanılarak sınıflandırma işlemi gerçekleştirilmiştir. Öznitelik çıkarımında tını özellikleri dikkate alınmıştır. Veri seti olarak GTZAN kullanılmıştır. En yüksek başarımlar SVM ile %88.9 olarak gözlemlenmiştir [13].

Thiruvengatanadhan'ın çalışmasında MFCCs özellikleri çıkarılarak pop, rock ve klasik olmak üzere 3 farklı türe müzik sınıflandırması yapılmıştır. SVM algoritması kullanılarak yapılan çalışmada polinomal çekirdek fonksiyonu için %88, gaussian çekirdek fonksiyonu için %91, sigmoidal çekirdek fonksiyonu için ise %87 başarımlar elde edilmiştir [14].

Şimşekli'nin çalışmasında Bas çizgisi tabanlı otomatik müzik sınıflandırması sistemi ele alınmıştır. Veri seti olarak 3 kök ve 3 yaprak yapısından oluşan McKay ve Fujinaga'nın MIDI verisi kullanılmıştır. Müzik türleri öncelikle 3 kategoriye daha sonra bu kategorilere ait 3'er alt kategoriye olmak üzere toplam 9 alt kategoriye sınıflandırılmıştır. Öznitelik çıkarımında "melodik aralık histogramları" kullanılmış olup ardışık 2 notanın uyumlarını baz alan PWED metriğinden yararlanılmıştır. Sınıflandırma algoritması olarak ise K-NN kullanılmıştır. Bu şekilde yaklaşık olarak yaprak düğümlerde %86.67 kök düğümlerde ise %100'e kadar doğruluk elde edilmiştir [15].

McKay ve Fujinaga'nın çalışmasında K-NN ile MIDI formatındaki müzik verilerinden önce 3 alt kategoriye (klasik, caz, popüler) sınıflandırma işlemi yapıp daha sonra bu alt kategorilerin her biri de 3'er alt kategoriye sınıflandırılmıştır. İlk 3 alt kategoriye (kök düğümler) sınıflandırma işleminde yaklaşık %98 doğruluk oranı yakalanmıştır. 2. alt kategoriye bölme işleminde ise (yaprak düğümler) doğruluk oranı %90'lara kadar düşmüştür [16].

Kotsifakos ve arkadaşlarının çalışmasında MIDI tipinde veriler ile süre ve perde bilgileri kullanılarak 4 farklı türden 100 parça için K-NN algoritması ile dört

kategoriye mzik sınıflandırması yapılmıřtır. K deęeri 4 seęildięi durumda %43.48 ile en yksek bařarıma ulařılmıřtır [17].

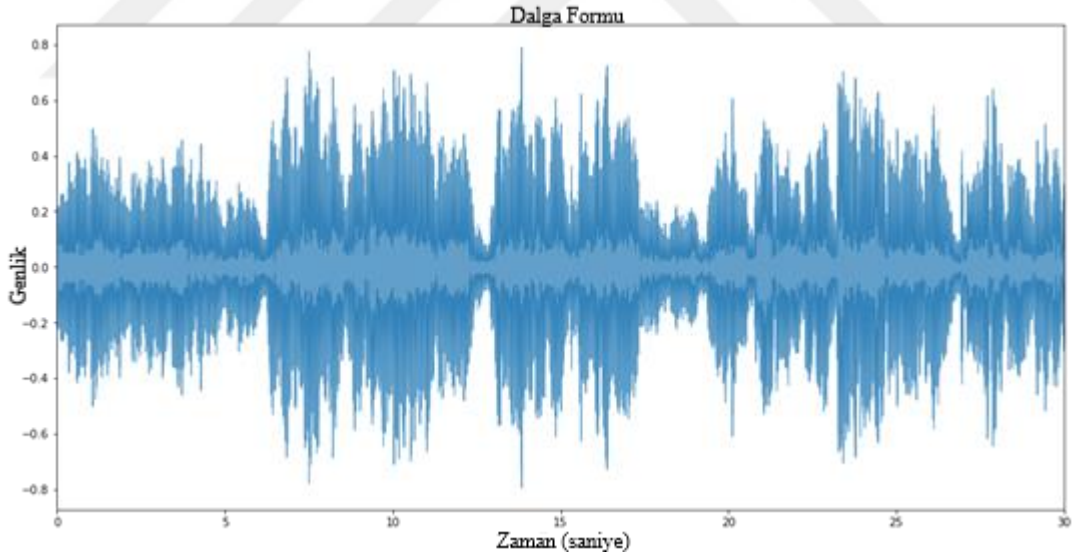


## 2. MÜZİKTE SES SİNYALLERİNİN İŞLENMESİ

Python programlama dilinde librosa kütüphanesi kullanılarak ses sinyallerinden çeşitli özellikler çıkarmak mümkündür. Bu tez çalışmasında da librosa kütüphanesinden yararlanılarak ses sinyalinin özellikleri elde edilmiştir. Librosa; müzik analizi ve ses analizi için geliştirilmiş olan bir python kütüphanesidir.

Ses sinyalleri işlenirken birtakım filtreleme, ön işleme ve çeşitli matematiksel işlemler ile ilgili özellikler elde edilmektedir. Temelinde birçok karmaşık matematiksel işlemi barındıran bu işlemler librosa kütüphanesi ile birkaç satır kod parçası ile rahatlıkla elde edilebilmektedir.

Aşağıdaki şekilde, kullanılan veri setine ait 30 saniyelik bir parçadan python programlama dilinde elde edilmiş bir dalga formunun (waweform) genlik-zaman grafiği verilmiştir (Şekil 2.1).



Şekil 2.1 : Veri setine ait bir parçanın python ile elde edilmiş genlik-zaman grafiği.

### 2.1 Örneklem Oranı (Sample Rate)

Örneklem oranı; bir saniyede alınan ses örneklerinin sayısını ifade eder. Sayısal değeri hertz cinsinden ölçülür. Bir ses örneği, yalnızca belirli bir zaman noktasında ölçülen akustik dalga değerini temsil eden bir sayıdır. Örneğin, örnekleme oranı 5000 Hertz olduğu varsayıldığı durumda, örneklerin tam olarak 1/5000 saniye arayla alınmaları

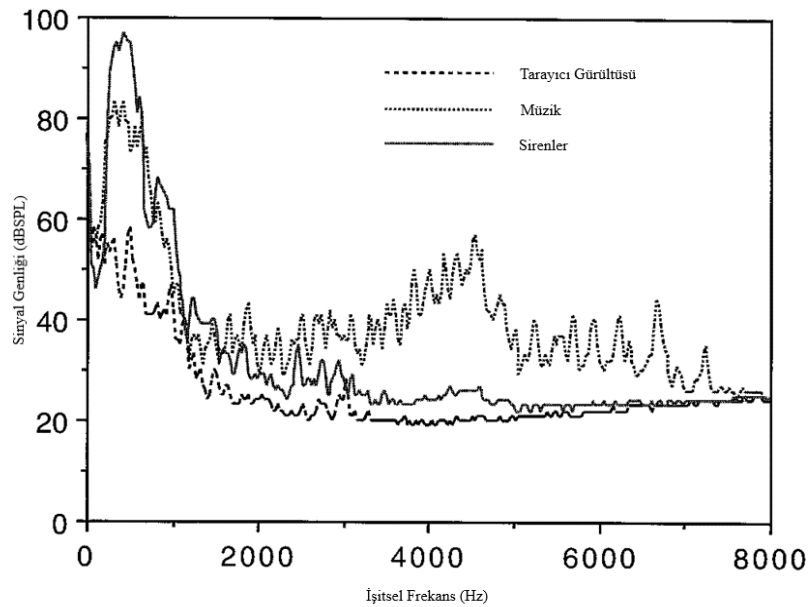
gerekir. Bu durumda 1/5000 sayısı örnekleme aralıdır. Örnekleme oranı ise bu deęerin çarpımsal tersidir. Bu durum; videolardaki ardışık görüntüleri, insan beyninin kesintisiz bir akış gibi algılaması gibi düşünülebilir. Örnekleme oranı, ses dosyalarının boyutlarını da doğrudan etkileyen bir unsurdur.

Müzik dosyaları için en sık kullanılan örnekleme oranları 22.05 kHz veya 44.1 kHz'dir. Yapılan çalışmada kullanılan parçaların örnekleme oranları da 22.05 kHz olarak alınmıştır. Bu deęer arttıkça verilerin bellekte kapladığı yer miktarı da orantılı olarak artmaktadır. Bu frekans deęeri saniyede alınan örnekleme sayısını ifade ettiğinden ötürü ses sinyalinden elde edilecek özellikler için bu oran önem arz etmektedir.

## 2.2 Frekans ve Genlik

Frekans, titreşimin hızıdır. Bu hız sesin perdesini belirler. Bir saniyede meydana gelen dalga devirlerinin sayısı olarak ölçülür. Güçlü bir düzenli dalga biçiminin olduğu müzikal sesler için anlamlıdır. Frekansın birimi Hertz'dir. Genlik ise titreşimin boyutunu ifade eder. Sesin ne kadar gürültülü olduğunu belirler. Daha fazla titreşimler daha yüksek sesler çıkarır. Sesin şiddetini kontrol etmede veya dengelemede genlik önemlidir [18].

Aşağıdaki şekilde; tarayıcı gürültüsü, müzik sesi ve siren sesleri olmak üzere 3 farklı sese ait seslerin frekans-genlik grafikleri verilmiştir (Şekil 2.2).



Şekil 2.2 : Çeşitli ses sinyallerine ait frekans-genlik grafikleri [18].

### **3. MAKİNE ÖĞRENMESİ ALGORİTMALARI VE SINIFLANDIRMA PROBLEMİ**

Makine öğrenmesi algoritmaları sınıflandırma problemlerinde yaygın olarak kullanılmaktadır. Bu bölümde sınıflandırma problemine bir bakış açısı sunulacak ve sınıflandırma kavramına ait bazı tanımlar verilecektir. Ayrıca yapılan tez çalışmasında kullanılan algoritmalarından bahsedilecektir.

#### **3.1 Makine Öğrenmesi**

Makine öğrenmesi bilgisayarın nasıl öğrenebileceği ile ilgilenen bir araştırma alanıdır. Makine öğrenmesi ve yapay zekâ kavramlarının temelleri 1950’lerde Alan Turing tarafından öne sürülen “düşünebilen makineler”, “kendi kendine öğrenebilen makineler” kavramları ile atılmıştır. Makine öğrenmesi, yapay zekânın istatistik ve bilgisayar bilimlerinin birleştiği bir alt dalıdır [19].

Makine Öğrenmesi, veri topluluklarının iç doğasının kavranmasına dayanan ve bunu da çeşitli algoritmalarla gerçekleştiren bir disiplindir. Sistem, karmaşık verileri tanıyarak anlamlandırıp bu verilerden tutarlı ve anlamlı sonuçlar çıkarmaktadır. Özünde sistem tarafından bir modelleme yapılmaktadır [19,20].

Makine öğrenmesinde modelleme yaparken hiperparametrelerin doğru seçilmesi oldukça önemlidir. Hiperparametreler, modelin parametreleri nasıl seçileceğini belirler. Daha yüksek başarımlar için bazı sınıflandırma problemlerinde hiperparametre grupları kullanılabilir.

Makine öğrenmesi; veri bilimi, derin öğrenme ve yapay zekâ gibi çalışma alanları ile yakından ilgilidir. Makine öğrenme algoritmaları; denetimli öğrenme, denetimsiz öğrenme, yarı denetimli öğrenme ve aktif öğrenme olmak üzere 4 ana başlıkta incelenmektedir. Günümüzde oldukça popüler hale gelen makine öğrenmesi; tıp, sanat, biyoloji gibi disiplinlerde de kullanılmaktadır.

### **3.1.1 Denetimli öğrenme (Supervised learning)**

Denetimli öğrenmede, verilerin içeriği ve bu verilerden istediğimiz çıktıların ne olması gerektiği bilinmektedir. Veriler etiketlidir. Denetimli öğrenmede veriler arasındaki ilişkinin öğrenilebilmesi için, verileri ve o verilerden çıkan sonuçları makineye tekrar baştan verip bu bilgilerden bir fonksiyon üretilir. Harici verileri kullanarak genel bir model üretir [21]. Denetimli öğrenmede “regresyon” ve “sınıflandırma” olmak üzere iki problem vardır. Regresyon probleminde, girdi değişkenleri sürekli fonksiyonlara eşlemeye çalışılır. Sınıflandırma probleminde ise sonuçlar ayrı ayrı çıktılar için tahmin etmeye çalışılır.

En popüler denetimli öğrenme algoritmaları başlıca K-NN, karar ağaçları, SVM, naive bayes, lojistik regresyon, rassal ormandır.

### **3.1.2 Denetimsiz öğrenme (Unsupervised learning)**

Denetimsiz öğrenme, verilerin etiketsiz olduğu ham verilerden anlamlı veriler çıkararak bir makine öğrenmesi çeşididir. Bu modelde etiketli veriler olmadığından doğru cevaplar da yoktur. Giriş verilerini yorumlayıp işleyerek bir kümeleme işlemi yapar. Girişte X ve Y kümelerine ait veriler olduğu varsayıldığı durumda bunların herhangi bir etiket bilgisi olmaksızın bu algoritma X ve Y'yi ayrı ayrı kümeleyip sonucu gösterecektir. Denetimsiz öğrenme “öğretici olmaksızın” öğrenme olarak tanımlanabilir [22].

### **3.1.3 Yarı denetimli öğrenme (Semi supervised learning)**

Yarı gözetimli öğrenme, eğitim aşamasında etiketlenmiş ve etiketlenmemiş verilerin her ikisinin de kullanımına imkân sağlayan bir öğrenme türüdür. Denetimli ve denetimsiz öğrenmenin algoritmalarının özelliklerini birlikte taşır. Yarı denetimli öğrenmenin temel amacı, denetimsiz öğrenme ve denetimli öğrenmenin dezavantajlarını ortadan kaldırabilmektir [23].

Bu algoritmada bir miktar etiketli veriyle etiketsiz verilerin bir arada tutularak çalışılması sonucu başarımın, denetimli ve denetimsiz öğrenmeye göre daha yüksek tutulması hedeflenir. Algoritma kullanılan verilerin etiketli olup olmama duruma göre seçim yapmaktadır.

### 3.1.4 Aktif öğrenme (Active learning)

Aktif öğrenme, bir modeli eğitmek için gereken etiketli veri miktarını oldukça yüksek oranda düşürebilen bir tekniktir. Bu teknikte, tüm veriler için etiketleri bir kerede almayıp, modelin en çok karıştırdığı verilere öncelik verir ve sadece bunlar için etiket ister. Bu işlem performansı ve verilerin verimliliğini en üst seviyeye çıkarabilmek için yapılır [24]. Model daha sonra bu az miktarda etiketlenmiş veri üzerinden eğitilip tekrardan bir miktar daha etiket ister. Model, en kritik örneklere öncelik vererek, en yararlı bilgileri sağlamayı amaçlar. Sonuç olarak hem zamandan tasarruf edilmiş olup hem de gereksiz etiketlemelerin önüne geçilip başarılı modeller elde edilebilir.

### 3.1.5 Aşırı öğrenme (Overfitting)

Makine öğrenmesinde sıkça rastlanılan temel problemlerden bir tanesi aşırı öğrenme durumudur. Makine öğreniminde bir sınıflandırma probleminde eğitim veri seti ve test veri setinden, eğitim veri setine ait verilerin bilgilerinin ezberlenmesi ve yeni veriler geldiği zaman modelin ne yapacağını bilememesinden ötürü tutarsız sonuçlar üretmesi durumu “aşırı öğrenme” olarak tanımlanmaktadır [25].

Sistem ezberleme yaptığı için verilerin özelliklerini ezberlediği için artık küçük bir değişiklikte dahi istenenden çok uzak sonuçlar vermeye meyilli olacaktır. Bu problemin önüne geçmek için veri setini mümkün olduğunca genişletmek ilk akla gelen çözümlerdendir. Etkisi pek fazla olmayan özelliklerin etkisini sınırlamak da aşırı uyumu engellemek için uygulanabilmektedir [26].

### 3.1.5 Sınıflandırma

Makine öğrenmesinde sınıflandırma, bir girdi için belirli bir sınıfın tahmin edildiği bir modelleme problemini ifade eder. Nihâi olarak çıktıların bir kısmı pozitif kalan kısmı ise negatif değerlere ayrılmış durumda olur [27].

Sınıflandırmada yalnızca 1 veya 0 değerlerinin üretildiği yalnızca iki sınıfın var olduğu bir durum vardır. Bunun dışında çok sınıflı, ikiden fazla türe sınıflandırma durumu vardır. Sınıflandırma algoritmalarında 2 temel adım vardır. Bunlar: eğitim aşaması ve test aşamasıdır.

Eğitim aşaması: Sınıflandırma işlemini gerçekleştirmek için yapılan öğrenme aşamasıdır. Eğitim örnekleri kullanılarak oluşturulur.

Test aşaması: Modelin kullanıldığı sınıflandırma aşamasıdır. Eğitim aşamasındaki öğrenilen durumlardan ilgili model oluşturulup uygulanarak sınıflandırma sonuçları elde edilir.

Çeşitli sınıflandırma algoritmalarının her birinin birbirine göre avantajları ve dezavantajları vardır. Başarılı sonuçlar alabilmek için veri setinin niteliğine göre sınıflandırıcı seçimi yapmak gerekmektedir. Sınıflandırma yaparken; verilerin türü, boyutu, işlenebilirliği dikkate alınması gereken önemli unsurlardır.

Yalnızca tek sınıflandırıcıyı kullanmak veya birden fazla sınıflandırıcıyı eğitip ardından tüm sınıflandırıcıların sonuçlarına dayalı olarak bir sınıflandırma kararı vermek de mümkündür. Sınıflandırma problemlerini çözmek hem pratik hem de teorik olarak ilgi görmektedir [28]. Sınıflandırma algoritmalarından başlıcaları; K-NN, rassal orman, naive bayes, lojistik regresyon, SVM, karar ağaçları, yapay sinir ağlarıdır.

### **3.1.6 Karmaşıklık matrisleri**

Beklenmedik durum tabloları olarak da adlandırılan karmaşıklık matrisleri, değişken çağırma algoritmaları gibi sınıflandırıcıları değerlendirmek için kullanılır. İkiye iki karmaşıklık matrisleri, bir algoritma tarafından atanan değişken etiketleri ile beyan edilmiş bir doğruluk kümesinden gelen etiketler arasındaki ilişkiyi sunar. İkiye iki karmaşıklık matrisindeki dört temel değer: gerçek pozitif (TP), gerçek negatif (TN), yanlış pozitif (FP) ve yanlış negatif (FN), algoritma performansını değerlendirmek için kullanılır [29].

Gerçek pozitif, modelin pozitif olan sınıfı doğru bir şekilde pozitif olarak öngördüğü sonuçtur. Benzer şekilde, gerçek negatif, modelin negatif sınıfı doğru tahmin ettiği sonuçtur. Yanlış pozitif, modelin pozitif sınıfı yanlış tahmin ettiği sonuçtur. Yine benzer şekilde, yanlış negatif modelin negatif sınıfı yanlış tahmin ettiği sonuçtur.

Aşağıda verilen şekilde doğru tahminlerin ve yanlış tahminlerin karmaşıklık matrisinde nasıl konumlandığı gösterilmiştir (Şekil 3.1).



**Şekil 3.1 :** Karmaşıklık matrisi.

Şekilde görüldüğü üzere karmaşıklık matrisinde başarıya pozitif yönde etki sağlayan doğru tahminler sol üst köşeden, sağ alt köşeye çizilen köşegen ile temsil edilmektedir.

### 3.1.7 Sınıflandırma başarımları ile ilgili metrikler

Sınıflandırma problemlerinde doğru model seçiminde bakılacak metrikler önem arz etmektedir. Çeşitli ölçümsel metrikler mevcuttur. Bunlardan başlıcaları; doğruluk (accuracy), kesinlik (precision), duyarlılık (recall) ve F1 skoru (F1 score) olarak ön plana çıkmaktadır.

Doğruluk, modelin başarımını ölçmek için kullanılan bir metriktir. Formülü aşağıdaki gibidir.

$$\frac{TP+TN}{TP+TN+FP+FN} \quad (3.1)$$

Kesinlik, pozitif tahmin edilen değerlerin, gerçekte ne kadarının pozitif olduğunu gösteren bir metriktir. Matematiksel formülü aşağıdaki gibidir.

$$\frac{TP}{TP+FP} \quad (3.2)$$

Duyarlılık, pozitif tahmin edilmesi gereken durumların ne kadarının pozitif olarak tahmin edildiğini gösteren bir metriktir. Matematiksel formülü aşağıdaki gibidir.

$$\frac{TP}{TP+FN} \quad (3.3)$$

F1 skoru, kesinlik ve duyarlılık değerlerinin harmonik ortalamasını gösteren bir metriktir. Bazı durumlarda doğruluk metriği yerine F1 skurun tercih edilmesinin en önemli sebebi eşit dağılmayan veri topluluklarında yapılabilecek hatalı model seçiminin önüne geçmektir.

$$F1=2*\frac{\text{precision*recall}}{\text{precision+recall}} \quad (3.4)$$

### 3.2 K-NN Algoritması

Denetimli bir makine öğrenme algoritması olan K-NN algoritması ilk olarak Cover ve Hart tarafından önerilmiştir. Sınıflandırma problemlerinde sıklıkla kullanılan kolay uygulanabilirliği ile ön plana çıkan bir algoritmadır. Algoritma, test verileri ile eğitim verilerini karşılaştırarak öğrenmeye dayanır [30]. K-NN sınıflandırma ve regresyon problemlerinin her ikisi için de çözümler sunabilmektedir. Algoritma, parametrik olmayan bir sınıflandırma yöntemidir. Basit bir algoritma olmasına karşın birçok sınıflandırma problemine etkili çözümler sunmaktadır [31]. K-NN’de birbirine yakın elemanların benzer elemanlar olduğu mantığıyla davranır. Belirli bir ölçüye göre birbirine yakın olan elemanlar belli bir sınıfa aittir.

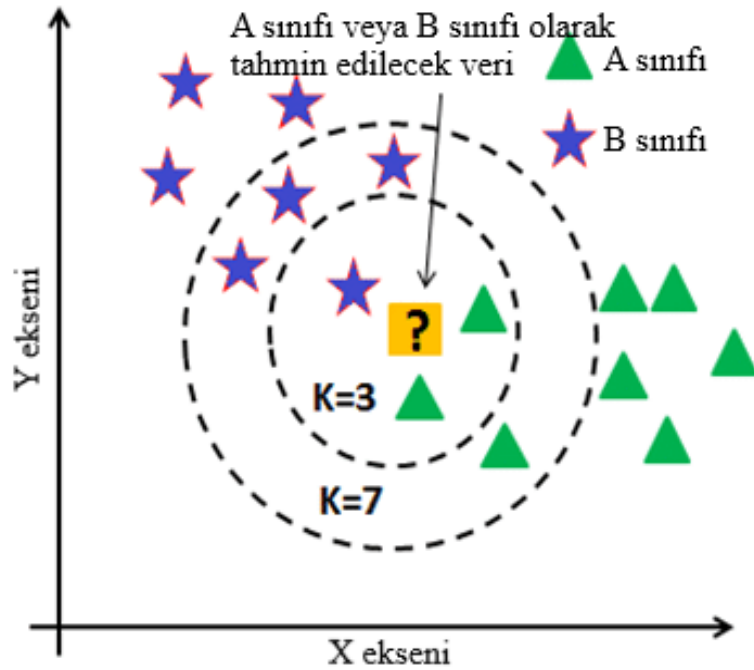
Algoritmada belirlenen K değerine göre bakılacak eleman sayısı belirlenir. Her yeni değer için K değerine göre belirlenen mesafe hesaplama yöntemine göre ilgili mesafe hesaplanır. Yeni değer için en yakın komşular tespit edilir. K-NN’de eklenecek olan yeni veri daha önceden belirlenen K değerine göre belirlenen mesafe hesaplama yöntemi ile ilgili sınıfa dahil edilir. Yeni bir veri ile karşılaşıldığında bu şekilde en uygun görülen sınıfa ataması yapılır.

K-NN tembel bir öğrenme algoritmasıdır. Çünkü eğitim verilerinden gerekli fonksiyonu öğrenmeyip, eğitim veri setini ezberler. Bu şekilde sınıflandırma işlemini gerçekleştirir. Tembel olmayan öğrenme algoritmalarında ise K-NN’nin aksine bir model uydurma veya eğitim adımı vardır.

K-NN algoritmasında K sayısını belirlemek oldukça önemli noktalardan bir tanesidir. K-NN'yi uygulamak için uygun bir k değeri seçmek gerekmektedir. Sınıflandırma başarısı belirlenecek olan K değerine oldukça bağlıdır. Bu yüzden K-NN algoritması için k tarafından önyargılıdır denebilir. K-NN'de N veri kaydının sınıflandırılması için, en yakın K adet komşu alınır ve bu bir N komşuluğu oluşturur. Çoğunluk oylaması yapıp mesafeye dayalı ağırlıklandırma dikkate alınarak veya dikkate alınmadan N için sınıflandırmaya karar vermek için kullanılır [32].

K sayısını belirlerken optimum sonuca yakın değerler elde edebilmek için genellikle 5,6,7 gibi değerler belirlenir. K değerinin çok düşük veya çok yüksek tutulması durumunda çoğu durumda sonuçlar istenen değerlerden uzak çıkabilmektedir. K değerini çok düşük belirlenmesi durumunda aşırı uyum durumu söz konusu olabileceken çok yüksek bir değer belirlendiği durumda ise elde edilecek olan başarımlar düşük seviyelerde kalabilecektir. K değerini belirlemek için özel bir yöntem olmadığından birkaç farklı K değeri ile sonucu görüp en uygun seçimi buna göre yapmak ideal bir tercih olacaktır.

K-NN algoritmasının; 2 farklı sınıfa ait verilerin, K değerinin sırayla 3 ve 7 olarak belirlendiği durumlar için verinin sınıfını belirlemesi (Şekil 3.2).



**Şekil 3.2** : K değerinin 3 ve 7 seçildiği bir örnek için K-NN algoritmasının K değerlerine göre en yakın komşulara yaklaşımı [33].

### 3.2.1 K-NN’de mesafe hesaplama yöntemleri

K-NN algoritmasında, komşu verilerin durumu baz alınarak gerekli mesafe hesaplamaları yapılır. Bu nedenle mesafe ölçümü sınıflandırma problemi için anahtar görevi görmektedir. Belirli bir noktanın en yakın komşuları daha önceden belirlenmiş olan K değerine göre tespit edilip mesafe ölçümleri yapılır. K-NN algoritmasında eğitim ve test verileri arasındaki mesafeyi hesaplamada sürekli değişkenler için çeşitli matematiksel yöntemler kullanır. Bunlardan en yaygın olarak kullanılanları; öklid uzaklığı, manhattan uzaklığı ve minkowski uzaklığı ve chebyshev uzaklığıdır.

Sınıflandırma problemlerinde çoğu zaman öklid mesafe fonksiyonu tercih edilmektedir. Bu mesafeler sadece sürekli değişkenler için geçerlidir. Sürekli olmayan değişkenler için ise farklı mesafe fonksiyonları kullanılmaktadır.

Öklid uzaklığı formülü;

$$d(x,y)=\sqrt{\sum_{i=1}^n (x_i-y_i)^2} \quad (3.5)$$

Buradaki x,y; hesaplanacak mesafeye ait noktaları temsil etmektedir.

Manhattan mesafesi formülü;

$$d(x,y)=\sum_{i=1}^n |x_i-y_i| \quad (3.6)$$

Buradaki x,y; hesaplanacak mesafeye ait noktaları temsil etmektedir.

Chebyshev mesafesi formülü;

$$d(x,y)=\max(|x_i-y_i|) \quad (3.7)$$

Buradaki x,y; hesaplanacak mesafeye ait noktaları temsil etmektedir. “max” ile x ve y arasındaki en büyük mesafe alınır.

### 3.3 SVM Algoritması

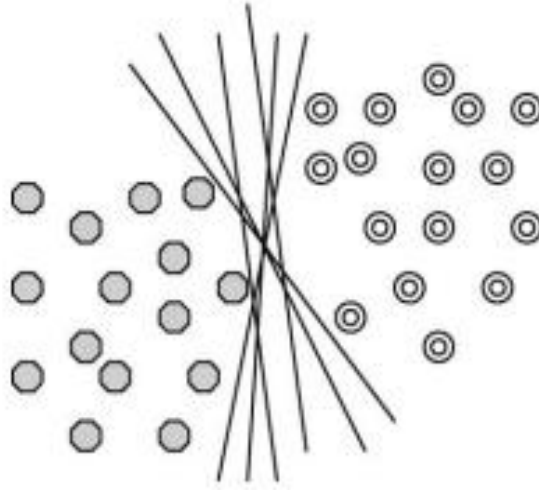
SVM (destek vektör makinesi) algoritması, regresyon ve sınıflandırma problemlerinin her ikisi için de kullanılabilir. SVM algoritmasının amacı verilerden en ilişkili örneklerin tespit edilip hiper düzlemler ile birbirinden ayırmaktır. SVM algoritması bahsi geçen ilişkili örneklerin hiper düzleme olan dik uzaklıklarını maksimize etmeyi amaçlar. Hiper düzlemler arasından en ideal ayrımı yapan optimum hiper düzlem olarak adlandırılır [34].

SVM algoritması, birçok sınıflandırma probleminde yaygın olarak kullanılmaktadır. SVM'nin müzik sınıflandırmasındaki başarımı, seçilecek olan parametrelere bağlı olmakla birlikte genel olarak diğer makine öğrenmesi algoritmalarına göre yüksektir. SVM'nin öğrenme aşamasındaki toplam işlem sayısı nispeten daha azdır. Ayrıca karmaşıklığı da daha düşüktür. Bu yüzden veriyi işleme performansı yüksektir. Özellikle büyük verilerde çalışmak için ideal bir algoritmadır. SVM'nin doğrusal ve doğrusal olmayan olmak üzere iki farklı türevi vardır. SVM, sınıflar arasında daha net ayrımlar yapabilmek için verileri daha fazla boyuttaki uzayda temsil etmek için doğrusal olmayan haritalama işlevini kullanır [35].

#### 3.3.1 Doğrusal SVM

Doğrusal SVM ile büyük veri kümelerinde kolayca çalışılabilmektedir. Doğrusal SVM iki farklı sınıfı birbirinden doğrular ile ayırmaya çalışır. Bu doğrular hiper düzlemlerdir. Hiper düzlemleri oluşturan noktalar ise destek vektörleridir. Eğitilmiş veri kümesine minimum mesafe sağlayan optimum hiper düzlemin bulunması amaçlanır [36]. Doğrusal SVM genellikle başarılı sonuçlar vermektedir. Ancak bazı durumlarda doğrusal olmayanlar kadar yüksek başarıma sahip olmayabilir.

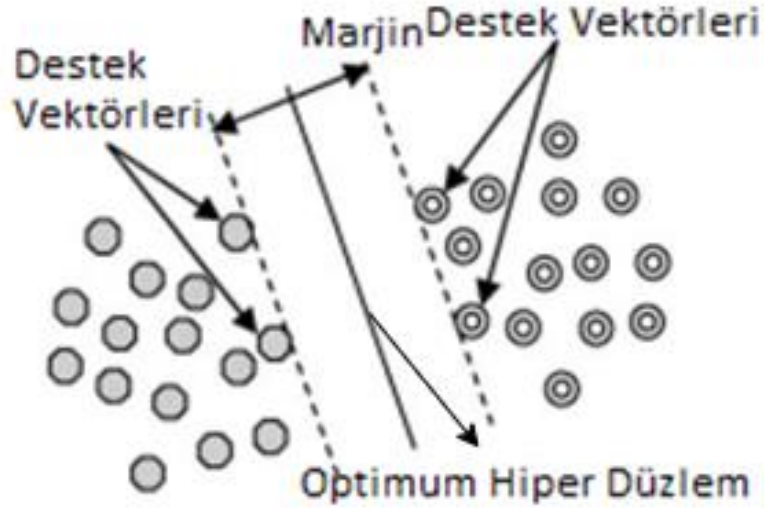
Aşağıdaki şekilde, 2 farklı sınıfa ait verilerin ayrıştırılması için oluşturulan hiper düzlemler gösterilmiştir (Şekil 3.3).



**Şekil 3.3 :** Doğrusal SVM'e ait çeşitli hiper düzlemler [36].

Hiper düzlemlerden, veri kümelerini en iyi şekilde ayırıştıranın seçilmesi gerekmektedir.

Aşağıdaki şekilde; destek vektörleri ve marjinden yola çıkarak elde edilen optimum hiper düzlem gösterilmiştir (Şekil 3.4).



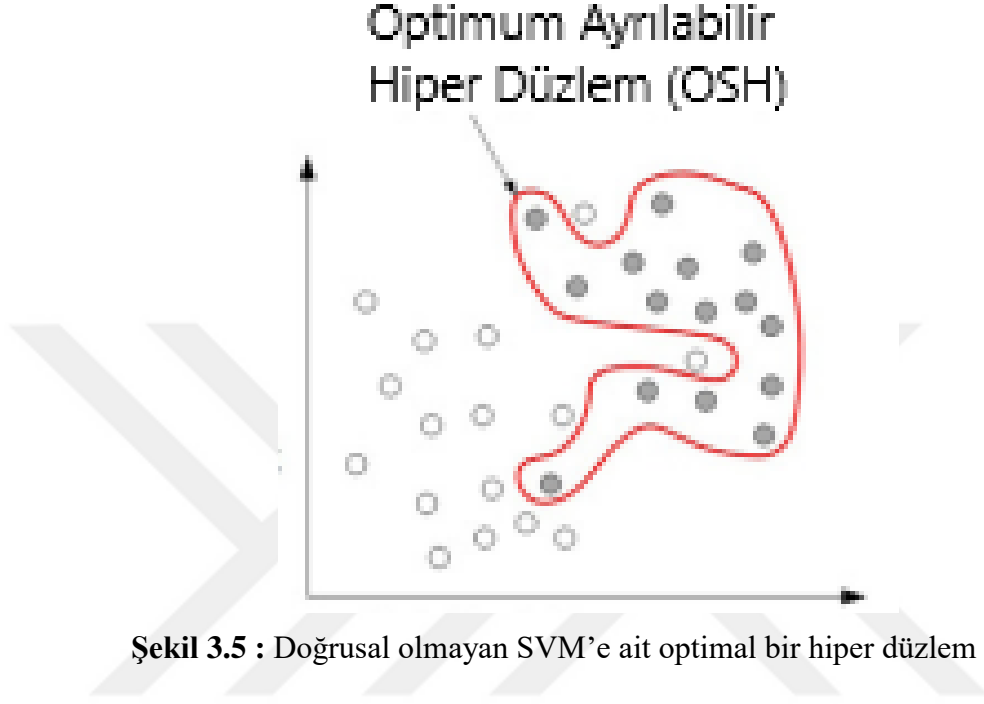
**Şekil 3.4 :** Doğrusal SVM'e ait optimum hiper düzlem [36].

### 3.3.2 Doğrusal olmayan SVM

Doğrusal olarak birbirinden ayıramayan sınıflar için bazı çekirdek yöntemleri vardır. Çok boyutlu sınıflandırmanın SVM algoritması ile uygulanabilmesi için çekirdek fonksiyonları kullanılır. Çekirdek fonksiyonları kullanılarak yüksek oranda doğruluk elde edilebilmektedir. Bu yöntemlerle doğrusal olarak ayıramayan sınıflar doğrusal olarak ayrılabilir hale getirilir ve bu şekilde sınıflandırma işlemi gerçekleştirilir.

Doğrusal olmayan SVM; doğrusal bir fonksiyonla ilgili verilerin ayrıştırılamadığı durumlarda kullanılmaktadır.

Doğrusal olmayan SVM ile verilerin birbirinden ayrıştırılması aşağıdaki şekilde verilmiştir (Şekil 3.5).



**Şekil 3.5 :** Doğrusal olmayan SVM'e ait optimal bir hiper düzlem [37].

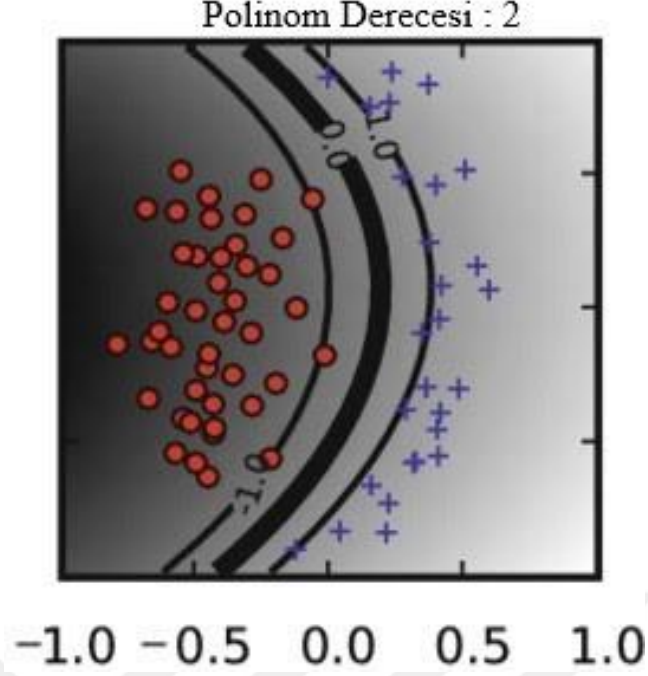
### 3.3.3 Çekirdek fonksiyonları

Çekirdek fonksiyonları, SVM'de tek boyutlu verilerin çok boyutlu olarak sınıflandırılmasına olanak tanıyan matematiksel fonksiyonlardır [38]. Çekirdek fonksiyonu sınıflandırma başarımına doğrudan etki etmektedir. Çekirdek fonksiyonlarından en yaygın olarak kullanılanlarından başlıcaları; polinomal çekirdek fonksiyonu, radyal tabanlı çekirdek fonksiyonu ve lineer çekirdek fonksiyonudur.

#### 3.3.3.1 Polinomal çekirdek fonksiyonu

Polinomal çekirdek, destek vektör makineleriyle ve diğer çekirdek modelleri ile birlikte kullanılabilen bir fonksiyondur. Bu fonksiyon, doğrusal olmayan modellerin öğrenilmesine izin verir.

Aşağıdaki şekilde polinomal çekirdek fonksiyonu ile polinom derecesi 2 olduğu durum için 2 farklı veri kümesinin birbirinden ayrıştırılması gösterilmiştir (Şekil 3.6).



**Şekil 3.6 :** 2. derece polinomal çekirdek fonksiyonunu ile verilerin ayrıştırılması [39].

Polinomal çekirdek fonksiyonunun matematiksel formülü aşağıdaki gibidir.

$$K(x_i, x_j, c, d) = (c + x_i^T x_j)^d \quad (3.8)$$

Formülde verilen  $d$  polinom derecesini ifade ederken,  $c$  ise  $c$  sabiti olarak adlandırılan bir sabit sayıdır.

### 3.3.3.2 Radyal tabanlı çekirdek fonksiyonu

Radyal tabanlı çekirdek fonksiyonu iki nokta arası benzerliği ve yakınlık durumlarını hesaplar. Matematiksel formülü aşağıda verilmiştir.

$$K(x_i, x_j, \sigma) = e^{-\frac{|x_i - x_j|^2}{2\sigma^2}} \quad (3.9)$$

Burada verilen  $\sigma$  sigma parametresidir. Doğruluk oranına önemli bir etkisi vardır. Formüldeki  $e$  ise Euler sayısını ifade etmektedir.

Literatürdeki uygulamalarda çoğunlukla daha yüksek başarıma sahip olduğu düşünülerek radyal tabanlı çekirdek fonksiyonunun kullanıldığı görülmektedir.

Kullanılacak veri seti çok büyük olmadığı durumlarda genellikle radyal tabanlı çekirdek fonksiyonu tercih edilmektedir.

### 3.3.3.3 Lineer çekirdek fonksiyonu

Veriler, doğrusal düzlemlerle birbirinden ayrıştırılır. Lineer çekirdek fonksiyonunun matematiksel formülü aşağıdaki gibidir.

$$K(x_i, x_j) = x_i^T x_j + c \quad (3.10)$$

Buradaki  $c$ ;  $c$  sabitini ifade etmektedir.

## 3.4 Lojistik Regresyon

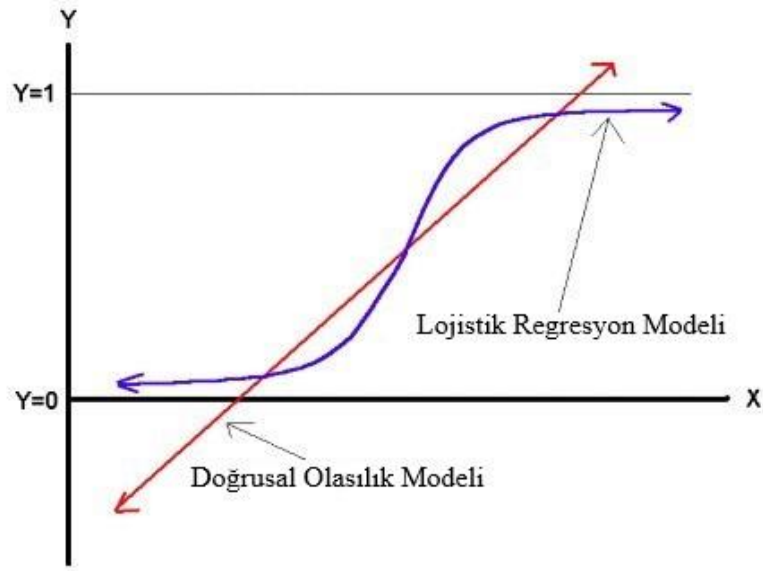
Lojistik regresyon alanındaki ilk çalışmalar 1944'lü yıllarda başlayarak Berkson tarafından yapılmış olup, 1972 yılında Finney, lojistik regresyonu probit analizine bir alternatif olarak önermiştir. Truett ve arkadaşları 1967 yılında, Halperin ve arkadaşları 1971 yılında lojistik regresyonun, ortalamasının 0 ve varyansın 1 olduğu normal dağılım varsayımları ihlal edildiğinde, diskriminant analizine alternatif olarak önermişlerdir [40]. Lojistik regresyon modelleri birçok alanda uygulanabilmektedir. Lojistik regresyon, tahmine dayalı analizler için kullanılmaktadır. Başlıca tıp, mühendislik, ekonomi ve biyoloji alanlarında olmak üzere birçok farklı alanda da uygulamaları mevcuttur.

Lojistik regresyon, bir veya birden fazla bağımsız değişken içeren veri kümesini analiz etmek için kullanılan istatistiksel bir yöntemdir. Doğrusal sınıflandırma problemlerinde kullanımı yaygındır.

Lojistik regresyon analizi, benzer analizlerin gerektirdiği varsayımlara ihtiyaç duymaksızın nispeten esnek bir biçimde model oluşturma imkânı sunması sayesinde tercih edilen çok değişkenli bir analiz yöntemi haline gelmiştir [41].

Lineer regresyona benzer bir yapıda olan lojistik regresyonun lineer regresyonla arasındaki temel farklılık; lojistik regresyonda bağımsız değişkenin kategorik olması lineer regresyonda ise bağımsız değişkenin sürekli olmasıdır.

Aşağıdaki şekilde, lojistik regresyon modeline ait bir eğri ve doğrusal olasılık modeline ait bir doğru gösterilmiştir (Şekil 3.7).

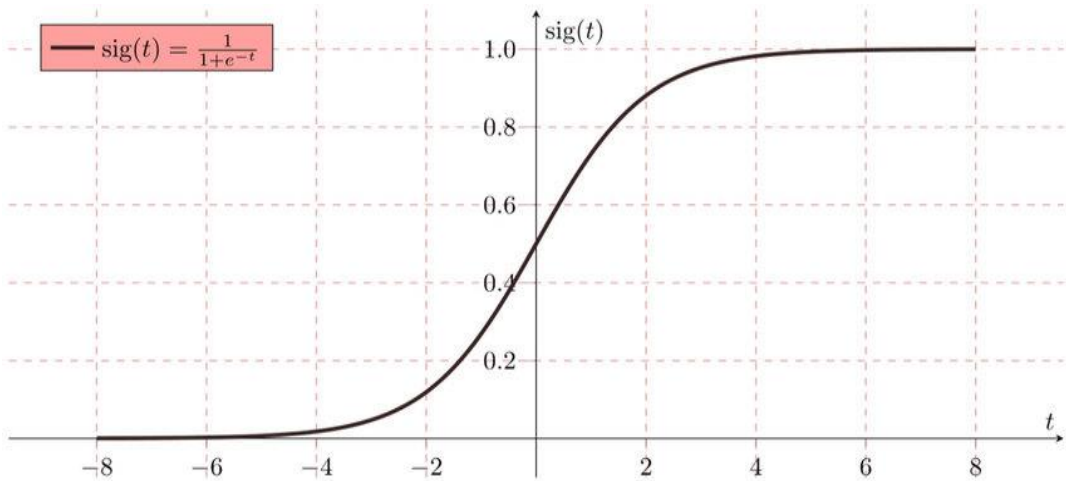


**Şekil 3.7 :** Lojistik regresyon ve lineer olasılık modellerine ait eğriler [42].

Lojistik regresyonda, lineer regresyondaki gibi verilerin arasına doğrudan düz bir çizgi çizilmeyip bir eğri uydurulur. Lineer regresyondan daha karmaşık bir maliyet fonksiyonu vardır. Bu fonksiyon genel olarak sigmoid fonksiyonudur.

Sigmoid fonksiyonu 0 ile 1 arasında sonuçlar verir. Sigmoid fonksiyonu öğrenmenin gerçekleştiği, makine öğreniminde yaygın olarak kullanılan bir aktivasyon fonksiyondur. Bu fonksiyon 0 ile 1 arasını sınırlama eğilimindedir. Dolayısıyla başarılı-başarısız veya pozitif-negatif gibi iki değere sahip bir sonucun tahmininde başarılı sonuçlar verebilir.

Aşağıdaki şekilde zamana bağlı olarak sigmoid fonksiyonunun 0 ile 1 arasındaki değişken değerleri görülmektedir (Şekil 3.8).



**Şekil 3.8 :** Sigmoid fonksiyonuna ait eğri [43].

Sigmoid fonksiyonunun matematiksel formülü,

$$\text{sig}(t) = \frac{1}{1+e^{-t}} \quad (3.11)$$

şeklindedir. Buradaki t zamanı ifade etmektedir.

Lojistik regresyonda sigmoid fonksiyonunun esnekliği ile sınıflandırma problemlerinde başarılı sonuçlar alınabilmektedir. Bu yöntem; kolaylıkla uygulanabilir olması dolayısıyla kolay yorumlanabilir olması ve sınıflandırma problemlerinde, seçilecek problemin türüne göre iyi sonuçlar vermesi sayesinde sınıflandırma problemleri için kullanılacak yöntemler arasındadır.

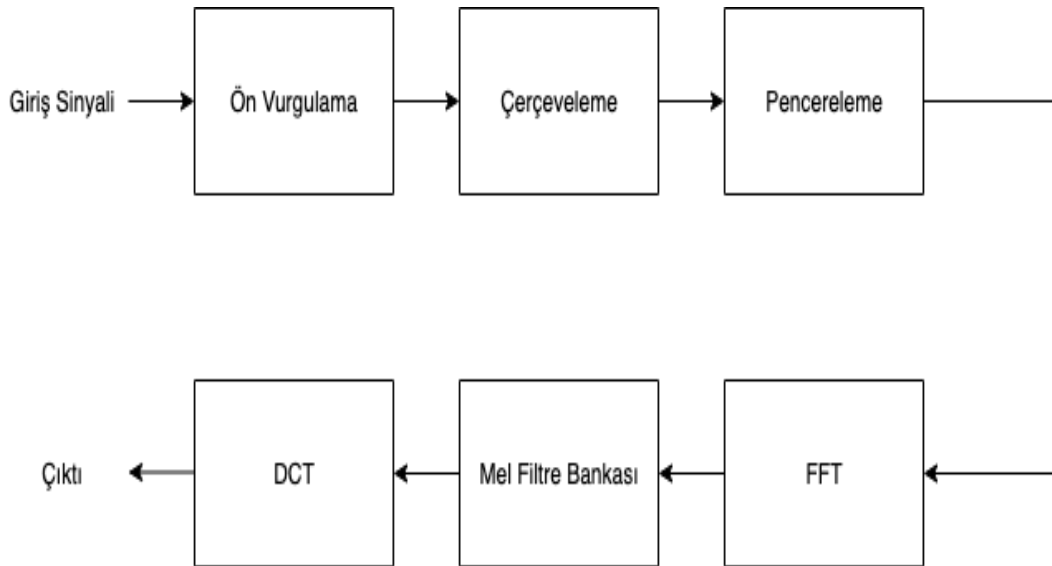


## 4. ÖZİNİTELİK ÇIKARIMI

Bu tez çalışmasında kullanılan özneliklerin tümü Python programlama dilinde Librosa kütüphanesi ile elde edilmiştir. Librosa kütüphanesi ile elde edilen öznelikler; Mel frekans kepstral katsayıları, sıfır geçiş oranı, izgesel merkez, izgesel bant genişliği, izgesel azalma, chroma STFT ve ortalama karesel hatadır.

### 4.1 Mel Frekans Kepstral Katsayıları (MFCCs)

Mel Frekans Kepstral Katsayıları (MFCCs), yaklaşık 30 yıl önce geliştirilmiştir. MFCCs; sesin kepstral temsilinden türetilirler. Birçok matematiksel hesaplamalar ve dönüşümler ile vektörel bir çıktı üretir. MFCCs ses tanımada başarılı sonuçlar göstermektedir [44]. Bu yüzden konuşma tanımada oldukça yaygın olarak kullanılmaktadır [44,45]. Aynı zamanda müzik bilgi erişiminde de tür sınıflandırmasında yüksek performansı ile dikkat çekmektedir [46]. MFCCs ses sinyalini 6 temel aşamada işler. Bunlar; ön vurgulama, çerçeveleme, pencereleme, FFT, Mel filtre bankası ve DCT'dir. Aşama aşama bu adımları içeren blok diyagramı Şekil 4.1'de verilmiştir.



Şekil 4.1 : MFCCs öznelik çıkarımı blok diyagramı.

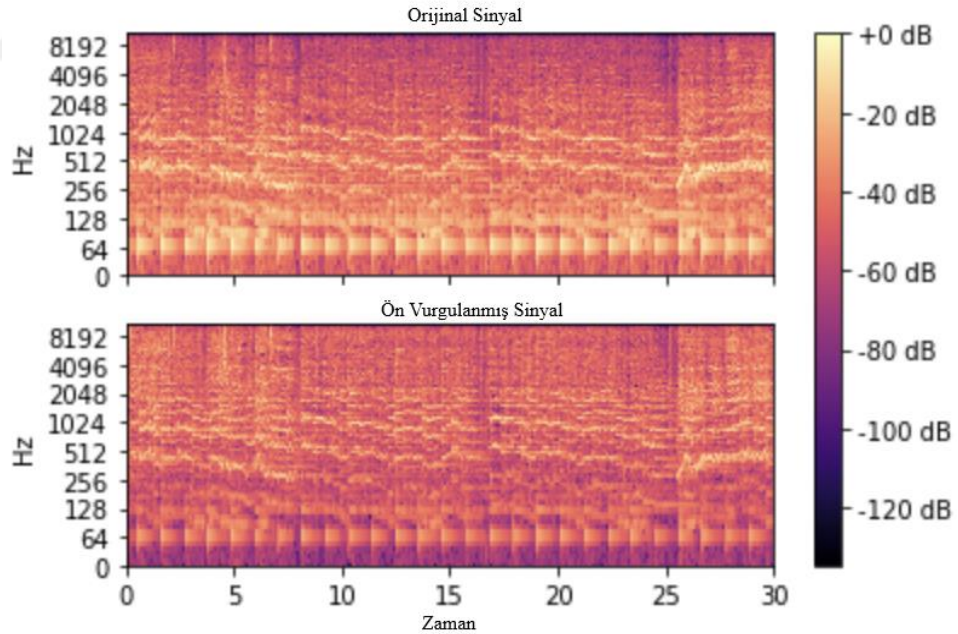
#### 4.1.1 Ön vurgulama

MFCCs’de ilk adım, yüksek frekansları güçlendirmek için sinyale bir ön vurgu filtresi uygulamaktır. Veri iletiminin çıkışındaki sinyalin kalitesini artırmak için ön vurgulama uygulanır. Sinyallerin yüksek veri hızlarında iletilmesinde, iletim ortamı bozulmalara neden olabilir, bu nedenle bu bozulmayı düzeltmek için iletilen sinyali bozmak için ön vurgu kullanılmaktadır. Bunun faydası, yüksek frekanslar genellikle düşük frekanslara kıyasla daha küçük büyüklüklere sahip olmasından ötürü frekans spektrumunu dengelemesidir [47]. Elde edilecek sinyal kalitesini artırmak için ön vurgulama uygulanır. Ön vurgu filtresi, aşağıda verilen denklemdeki gibi birinci dereceden filtre kullanılarak bir x sinyaline uygulanabilir:

$$y(t)=x(t)-\alpha x(t-1) \quad (4.1)$$

Formülde verilen  $\alpha$ , filtre katsayısını ifade ederken, t ise zamanı ifade etmektedir.

Aşağıda verilen şekilde, veri setine ait bir müzik parçasının Python programlama dili ile ön vurgulama sonucu elde edilmiş spektrogramı verilmiştir (Şekil 4.2).



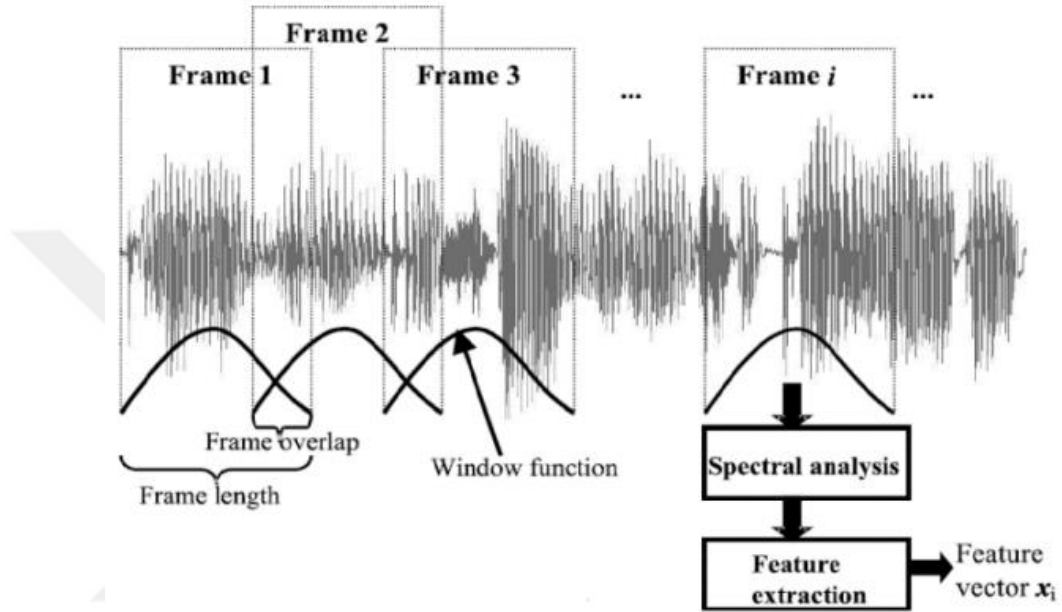
Şekil 4.2 : Ön vurgulama ile elde edilen spektrogram.

#### 4.1.2 Çerçeveleme

Konuşma işareti, parametrelerin sabit kaldığı kabul edildiği çerçeve olarak adlandırılan küçük parçalara ayrılmalıdır. Bu küçük parçaların FFT’si alınır ve her bir

çerçeveye örtüşme işlemi uygulanır. Çerçeveler örtüştürülürken %30-%75 değerleri arasında bir örtüşme oranı alınır. Tüm işaret boyunca FFT hesaplanmaması, izgesel bilgilerde kayıp yaşanmaması içindir. Her bir çerçeveye örtüşme uygulanır [48,49]. Örtüşme işlemleri sonrasında bir özellik vektörü elde edilir.

Aşağıdaki şekilde, çerçevelerin pencereleme fonksiyonu uygulanarak örtüştürülmesi ve öznitelik vektörünün çıkarılması durumu gösterilmiştir (Şekil4.3).

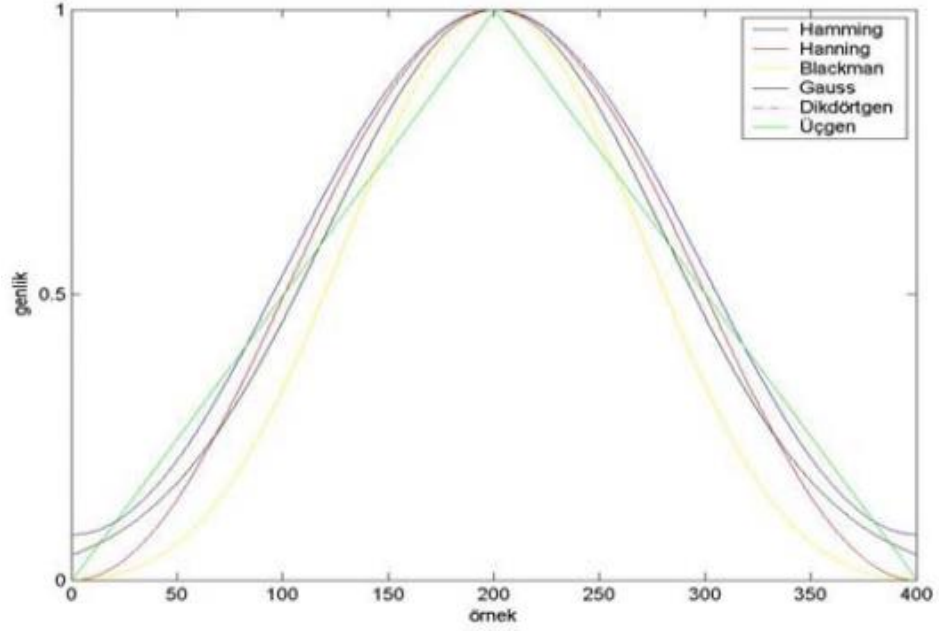


Şekil 4.3 : Çerçevelerin örtüştürülmesi [48].

#### 4.1.3 Pencereleme

Ses sinyallerinin, daha hızlı sonuçlar vermesi için bu sinyaller daha küçük parçalara ayrılarak daha kısa zamanlı özellikler incelenir. Bunun nedeni; daha fazla örneklem ile sinyallerden elde edilecek olan özelliklerin daha da artırılmasıdır. Bu şekilde çerçevelerdeki süreksizlik durumu ortadan kaldırılır. Fourier serisinin bir zaman alanı işlevi sınıfı kullanılarak kesilmesi olarak da bilinen pencereleme işlemini uygulamak için çeşitli fonksiyonlar mevcuttur. Bu fonksiyonlardan en fazla kullanılanı “hamming pencereleme fonksiyonu”dur. Bunun yanı sıra blackman, dikdörtgen, üçgen, Gauss ve Hanning gibi pencereleme fonksiyonları da mevcuttur. Pencereleme işlemi aynı zamanda bir tür sivrilme işlemidir [50]. Pencereleme işlevinde ses sinyalinin orta bölgeleri güçlendirilirken öte yandan kenar bölgeleri ise yumuşatılır.

Aşağıdaki şekilde 400 farklı örnek için 6 farklı pencereleme fonksiyonunun ürettiği eğriler verilmiştir (Şekil 4.4).



Şekil 4.4 : 400 örnek için çeşitli pencereleme fonksiyonlarına ait eğriler [48].

#### 4.1.4 Hızlı Fourier dönüşümü (FFT)

FFT, 1965'te tanıtılmış ve o tarihten bu yana oldukça yaygın bir kullanıma sahip olmuştur. Bunun sebebi FFT algoritması, DFT hesaplamaları için gereken aritmetik işlemlerin sayısını oldukça azaltıyor olmasıdır. Bu sayede, birçok soruna oldukça pratik çözümler sağlamıştır [51]. FFT; bir dizinin ayrık Fourier dönüşümünü (DFT) ya da ters ayrık dönüşümünü hesaplayan bir algoritmadır [52]. Sinyal; bulunduğu uzaydaki gösteriminden frekans uzayındaki gösterimine veya tersine dönüştürülür. DFT'de ise ayrık veri dizileri farklı frekans öğelerine ayrılır. Çoğu durumda kullanışlı olan bu durum, pratik bir seçenek olmaktan uzaktır. Bu yüzden DFT hesaplamaları için genellikle FFT algoritmaları kullanılmaktadır. FFT algoritmaları DFT dönüşüm matrisinin seyrek matrislere ayrıştırılması ile çalışır [53]. Bu şekilde DFT'nin karmaşıklığı

$$O(N^2) \quad (4.2)$$

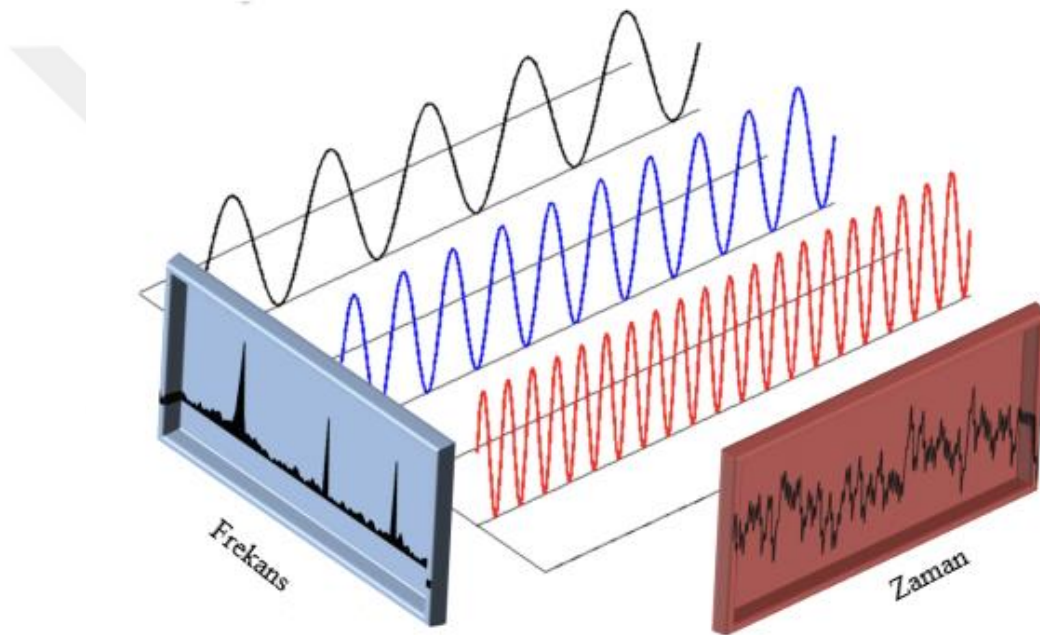
seviyesinden

$$O(N(\log N)) \quad (4.3)$$

seviyesine düşürülebilmektedir. Yukarıda verilen N değerleri verinin boyutunu ifade etmektedir.

Veri boyutunun çok yüksek olması durumunda FFT'nin çalışma hızı standart DFT'den çok daha fazladır. Birçok FFT algoritması mevcuttur. Bunlardan en popüler olanı Cooley–Tukey FFT algoritmasıdır. FFT algoritmaları müzikte de sıkça kullanılmaktadır [54,55]. Matematikçi Gilbert Strang, 1994 yılında FFT'yi "insan hayatındaki en önemli sayısal algoritmalarından biri" olarak tanımlamıştır. IEEE'nin Computing in Science & Engineering dergisi ise FFT'yi "20. Yüzyılın en önemli 10 algoritması" listesine dahil etmiştir.

Aşağıdaki şekilde, bir sinyale FFT uygulanması sonucu giriş sinyalinden elde edilen frekans-zaman grafiği verilmiştir (Şekil 4.5).

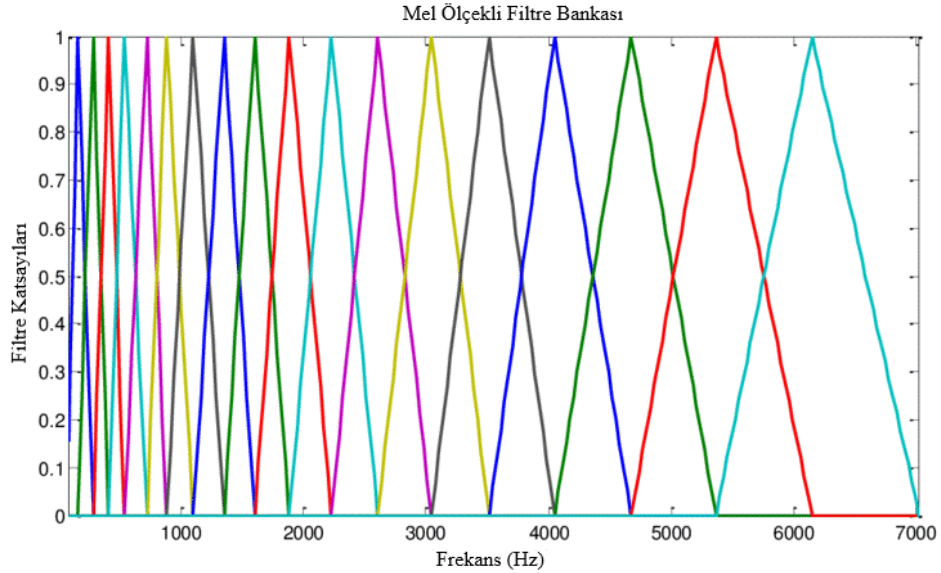


Şekil 4.5 : FFT'ye ait zaman-frekans grafiği [54].

#### 4.1.5 Mel filtre bankası

Mel filtre bankasını tanımlayan parametreler; Mel filtreleri, minimum frekans ve maksimum frekanstır. MFCCs özelliklerinin elde edilmesi aşamasındaki filtreleme işleminin yapıldığı Mel filtre bankasında, Fourier dönüşümünün bir modülü çıkarılır ve büyüklük spektrumu  $N$ 'ye  $P$  boyutunda bir matris olarak elde edilir. Büyüklük spektrumu, frekans çözünürlüğünü insan kulağının özelliklerine uyarlamak için Mel ölçeğine göre çarpıktır [56,57]. Ardından büyüklük spektrumu, tipik olarak, merkez frekansları tarafından tanımlanan bir dizi örtüşen üçgen filtreden oluşan Mel filtre bankası ile belirli bölümlere ayrılır [57].

Aşağıdaki şekilde; ilgili filtre katsayısı altında, Mel filtre bankası uygulanarak frekans değerleri gösterilmiştir (Şekil 4.6).



**Şekil 4.6 :** 20 Mel filtresi kullanılarak elde edilen Mel filtre bankası grafiksel gösterimi [58].

#### 4.1.6 Ayrık kosinüs dönüşümü (DCT)

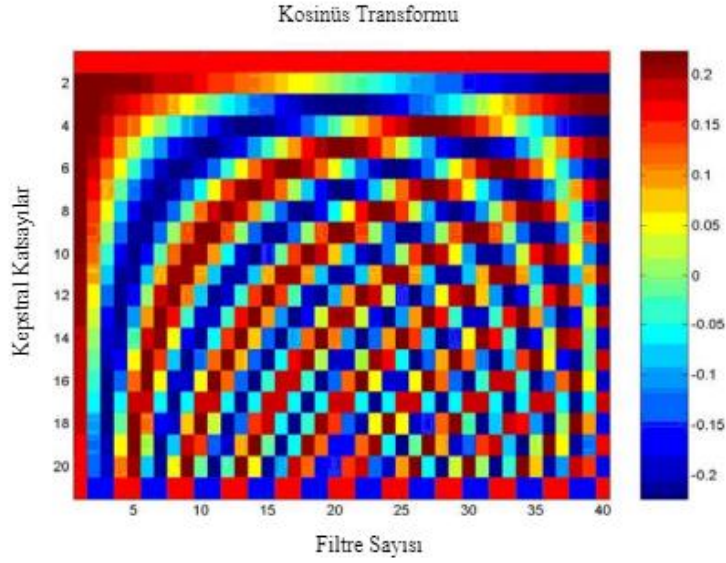
1972 yılında ilk olarak Hindistan kökenli bilim insanı Nasir Ahmed tarafından öne sürülmüş olan DCT, sinyal işlemede ve veri sıkıştırmada sıklıkla kullanılan bir dönüşüm tekniğidir. Sinyal işlemede oldukça yaygın olarak kullanılmaktadır.

DCT, bileşenleri kosinüs olan  $N$  gerçek tabanlı vektörleri kullanır. Bu temel vektörler ortogondur. Transpozesi ile tersi birbirine eşittir. FFT'den hızlı bir şekilde hesaplanırlar [59].

DCT, farklı frekanslarda salınan kosinüs fonksiyonlarının toplamı cinsinden sonlu veri noktaları dizisini ifade eder. DCT birçok alanda kullanılmaktadır. Sayısal sinyal işleme, kısmi diferansiyel denklemlerin çözümü için izgesel yöntemler, veri sıkıştırma, biyometrik tanıma sistemleri, kriptografi gibi alanlarda uygulanabilmektedir.

DCT, ayrık Fourier dönüşümüne benzer ancak yalnızca gerçek sayıları kullanan Fourier ile ilgili bir dönüşümdür. DCT'ler genellikle periyodik ve simetrik olarak uzatılmış bir dizinin Fourier serisi katsayıları ile ilgiliyken, DFT'ler yalnızca periyodik olarak uzatılmış dizilerin Fourier serisi katsayıları ile ilgilidir. DCT'ler kabaca iki katı uzunluktaki DFT'lere eşdeğerdir, gerçek veriler üzerinde eşit simetri ile çalışır.

Aşağıda verilen şekilde, bir DCT uygulamasına ait ilgili filtre katsayısı ve keprstral katsayılar altında DCT uygulanması sonucu renk ölçeğine göre matrisin aldığı değerler gösterilmiştir (Şekil 4.7).



Şekil 4.7 : Ayrık kosinüs dönüşümü ile elde edilen görsel [48].

#### 4.2 Sıfır Geçiş Oranı (Zero Crossing Rate)

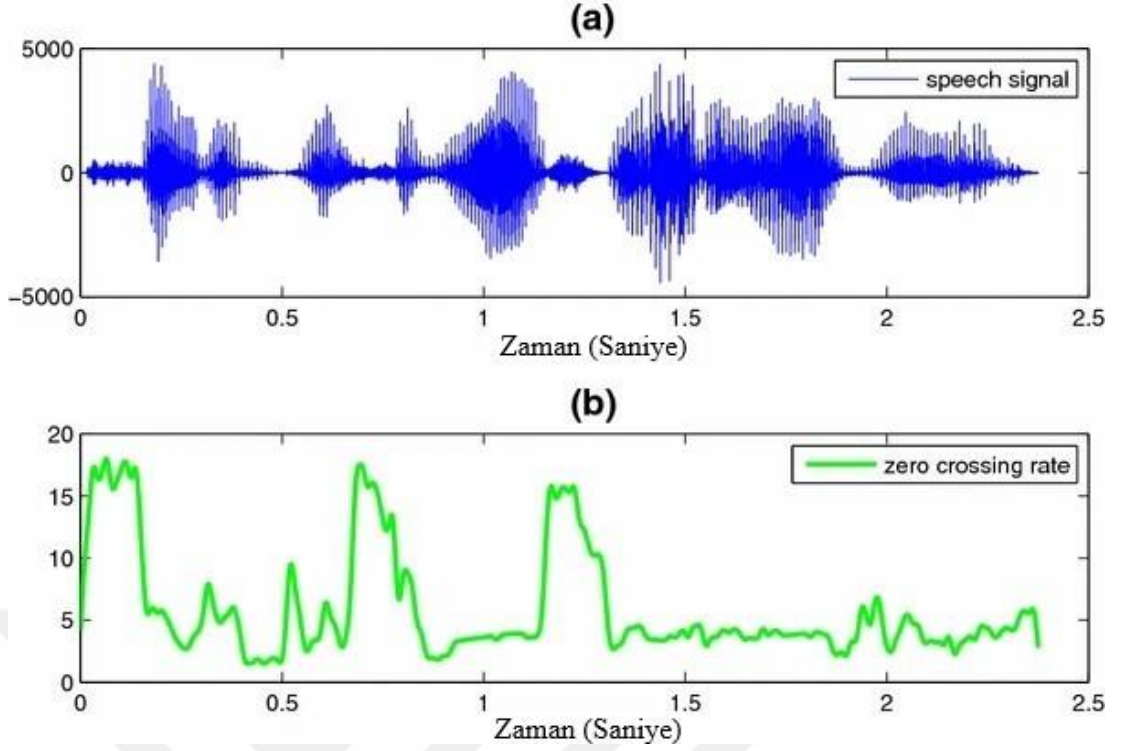
Ayrık zamanlı sinyaller bağlamında, farklı cebirsel işaretler varsa sıfır geçişin meydana geldiği ifade edilir. Sıfır geçişlerinin meydana gelme hızı, frekans içeriğinin bir ölçüsüdür [60].

Gürültüye bağlı sıfır geçişi azaltmak, sinyal üzerinde gürültünün etkisini düşürmek, daha belirgin kısımlarını elde etmek için bir eşik değeri dahil edilebilir. Sıfır Geçiş fonksiyonu matematiksel olarak aşağıdaki gibi ifade edilebilir;

$$zcr = \frac{1}{T-1} \sum_{t=1}^{T-1} 1_{R_{<0}}(s_t s_{t-1}) \quad (4.4)$$

Buradaki s, sinyal uzunluğunu göstermektedir. T ve  $1_{R_{<0}}$  ise işaretçi fonksiyondur.

Aşağıdaki şekilde, bir konuşma sinyalinden elde edilen sıfır geçiş oranının zamana bağlı grafiği verilmiştir (Şekil 4.8).



**Şekil 4.8** : Ses sinyalinden çıkarılmış sıfır geçiş oranı grafiği örneği [61].

Yüksek frekanslar yüksek sıfır geçiş oranlarını ve düşük frekanslar düşük sıfır geçiş oranlarını ifade etmektedir. Bu sebeple sıfır geçiş oranı ile frekans ve enerji dağılımı arasında güçlü bir korelasyon ilişkisi mevcuttur. Sıfır geçiş oranının yüksek olduğu durumda konuşma sinyalinin sesli olmadığı, sıfır geçiş oranının düşük olduğu durumda ise konuşma sinyalinin sesli olduğu yorumlanabilir [62].

Sıfır geçiş oranı vurmali çalgılarda daha yüksek ayırt ediciliğe sahiptir. Dolayısıyla özellikle vurmali çalgıları sınıflandırmada bu özellik kullanılabilir. Bunun yanı sıra ses tanıma uygulamalarında da kullanılmaktadır.

### 4.3 İzgesel Merkez (Spectral Centroid)

İzgesel merkez, sesin parlaklık ölçüsü ile ilişkilendirilmektedir. Bu ölçü, Fourier dönüşümü ile frekans ve büyüklük bilgisi kullanılarak kütle merkezinin hesaplanması ile elde edilir [63].

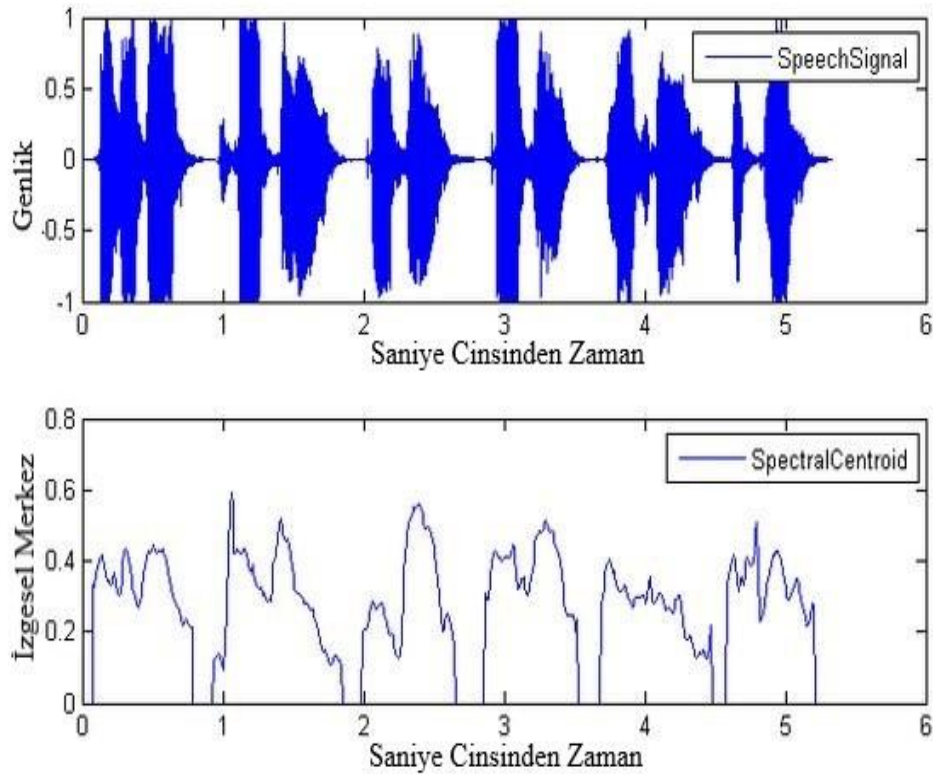
İzgesel merkez akustik enstrüman tonlarının en önemli algısal özelliklerinden bir tanesi olarak kabul edilmektedir. Düşük harmonilere göre daha yüksek harmonilerin gücünü ölçen izgesel merkez, bir tonun algısal parlaklığı ile güçlü bir şekilde ilişkilidir. Özellikle enstrümanların tonları hakkında fikir verici olabilmektedir [64].

Müzik sınıflandırmasında oldukça yaygın olarak kullanılan izgesel merkez, tınsal özelliklerden bir tanesidir. İzgesel merkez, bir spektrumun enerjisinin hangi frekansta merkezlendiğinin bir ölçüsüdür.

$$\frac{\sum_{n=0}^{n-1} f(n)x(n)}{\sum_{n=0}^{n-1} x(n)} \quad (4.5)$$

Verilen denklemdeki  $x(n)$  ağırlıklı frekans değerini ifade ederken  $f(n)$  ise merkez frekansını temsil etmektedir.

Aşağıdaki şekilde, bir konuşma sinyalinden elde edilmiş olan izgesel merkez değerleri verilmiştir (Şekil 4.9).



Şekil 4.9 : Sese ait sinyalden izgesel merkezin elde edilmesi [65].

#### 4.4 İzgesel Bant Genişliği (Spectral Bandwidth)

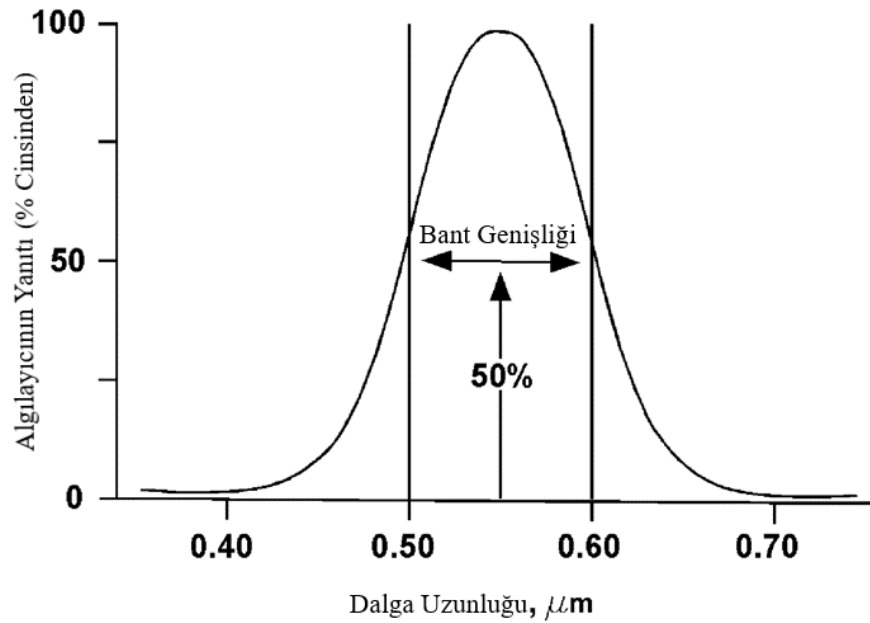
İzgesel bant genişliği, ses sinyaline ait izgesel özelliklerden bir tanesidir. Ses sinyalinin maksimum tepe noktasının yarısındaki bant genişliği olarak tanımlanmaktadır. Sesleri sınıflandırma problemlerinde sıklıkla kullanılan etkili bir tınsal özelliktir.

Spektral bant genişliğinin matematiksel formülü aşağıdaki gibidir.

$$\left( \sum_k S(k) (f(k)-f_c)^p \right)^{\frac{1}{p}} \quad (4.6)$$

Buradaki p değeri izgesel bant genişliğinin derecesini, S(k) ise izgesel büyüklüğü ifade etmektedir.

Aşağıdaki şekilde, spektral bant genişliğinin yüzdesel açıdan elde edilmesi gösterilmiştir (Şekil 4.10).

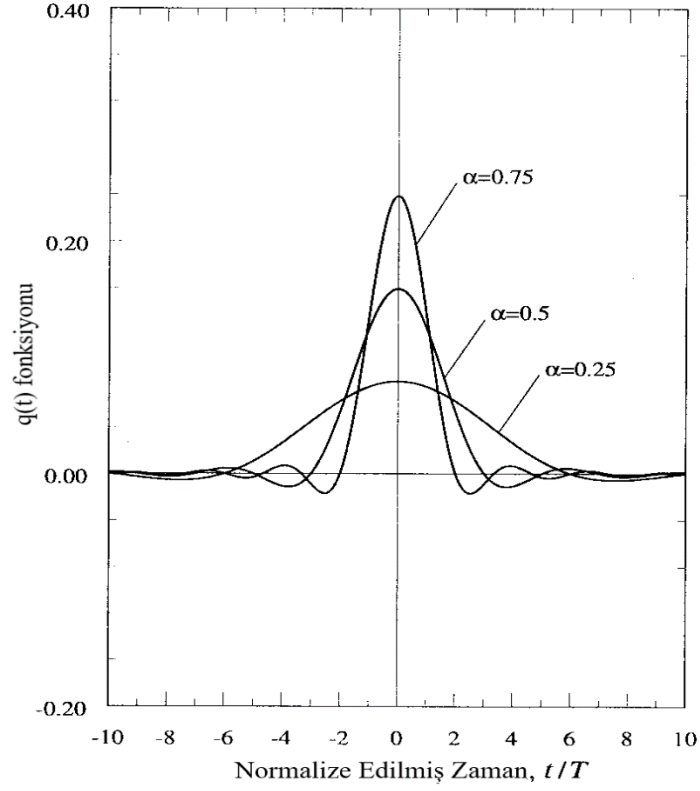


Şekil 4.10 : İzgesel bant genişliği grafiksel gösterimi [66].

#### 4.5 İzgesel Azalma (Spectral Rolloff)

İzgesel azalma, toplam izgesel enerjinin belirli bir yüzdesinin altında kaldığı frekansı ifade eder. Bu yüzde librosa kütüphanesinde varsayılan olarak %85'tir.

Aşağıdaki şekilde, çeşitli  $\alpha$  değerleri için  $q(t)$  fonksiyonuna göre izgesel azalmanın grafiği verilmiştir (Şekil 4.11).



Şekil 4.11 :  $q(t)$  fonksiyonu için izgesel azalmanın grafiği [67].

#### 4.6 Chroma STFT

Chroma tabanlı özellikler perde sınıfı profilleri olarak bilinen, müzik analizinde oldukça fazla kullanılan güçlü bir araçtır. Müziğe ait perde sınıflarını anlamlı bir şekilde kategorize edebilir [68]. Bir chroma vektörü, her bir perde sınıfının enerjisini gösteren 12 elemanlı bir özellik vektörüdür.

Akor tanımada oldukça yaygın olarak kullanılan chroma özelliği, sesin belirli bir zaman dilimindeki harmonik temsilindeki bir zaman serisinin chroma vektörlerini içerir [69].

Chroma STFT, sinyal işlemede kullanılan dönüşüm yöntemlerinden bir tanesidir. Örneklemeler üzerinde FFT uygulanarak elde edilen spektrumun “chromagram”da dikey bir eksenle görüntülenmesi sonucunda elde edilir [70].

#### 4.7 Ortalama Karesel Hata (RMSE)

Sayısal sinyal işlemede RMSE, bir sesin küresel enerjisidir. Ses örneklerinden her bir kare için ortalama karekök değerini ifade eder.

RMSE, aşağıdaki denklem ile hesaplanır.

$$x_{\text{rms}} = \sqrt{\frac{1}{n} \sum_{i=1}^n x_i^2} \quad (4.7)$$

Buradaki;  $x_i$ , i. karedeki sinyal genliğini, n ise örnek uzunluğundaki çerçevelerin sayısını ifade etmektedir [71].



## **5. BULGULAR VE YORUMLAR**

Bu bölümde yapılan tez çalışmasındaki sürecin genel akışı, yapılan çalışmanın sonuçları ve karşılaştırmalı analizleri sunulacaktır.

### **5.1 Veri Seti**

Bu çalışmada; webden elde edilen, 1. dakikadan 1:30. dakikaya kadar 30’ar saniyelik kesitleri kırılarak elde edilmiş olan, tamamı Türkçe müzik parçalarından oluşan wav uzantılı dosyaları içeren bir veri seti kullanılmıştır.

Wav uzantılı dosyaların boyutları sıkıştırılmış diğer formatlara göre oldukça yüksektir. Bu yüksek boyutlarda çalışmak oldukça zahmetli olduğundan kullanılan veri setinde her bir parçanın 30’ar saniyelik kesitlere bölünerek müzik parçalarından birer kesit alınarak işlenmiştir. Her bir parçadan alınan 30’ar saniyelik kesit yaklaşık olarak 5,5 megabayt boyutundadır. 1.00 ve 1.30. dakikalar arasının seçilmesinin nedeni yapılan uzun vadeli gözlemler sonucunda bu süre aralığının parçaların karakteristiğini genel olarak daha net bir şekilde yansıttığı düşünülmesidir. Bu tespit birçok parça üzerinde yapılan gözlemler sonucu elde edilmiştir.

Her bir sanatçıya ait en fazla 5 parça seçilerek çeşitlilik sağlanmış ve bu sayede mümkün olan en yüksek seviyede homojenlik sağlanmış olup elde edilecek başarımların ölçütlerinin daha sağlıklı bir şekilde gözlemlenmesi hedeflenmiştir. Veri setine ait sınıflandırılacak olan “Rock Müzik”, “Pop Müzik”, “Rap Müzik”, “Arabesk Müzik”, “Klasik Müzik”, “Dini Müzik” olmak üzere 6 temel tür belirlenmiştir. Veri seti her bir türe ait 100’er parça içermekle birlikte toplamda 600 parçadan oluşmaktadır.

### **5.2 Wav Müzik Formatı**

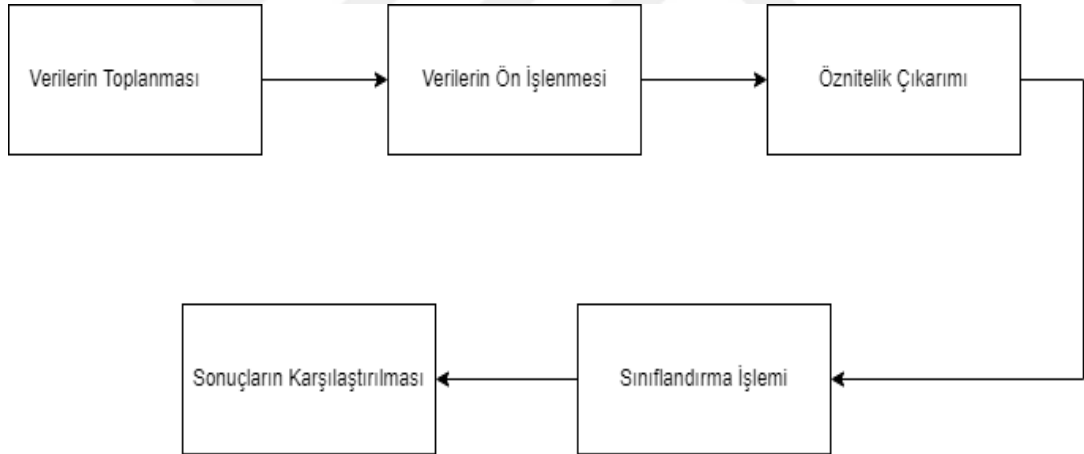
1991 yılında IBM ve Microsoft tarafından geliştirilip piyasaya sürülen wav müzik formatı sıkıştırılmamış bir ses dosyası formatıdır. Dolayısıyla kayıpsızdır. Yalnızca sayısallaştırılmış sinyalleri içerir. Kayıpsız olmasından ötürü mümkün olan en yüksek ses kalitesini sağlar. Bu nedenle wav, müzik endüstrisinde sıklıkla tercih edilen bir ses formatıdır. Wav formatının dezavantajı, sıkıştırılmamış veriler oldukları için çok fazla

depolama alanı kaplar. 3 dakikalık bir wav dosyasının boyutu yaklaşık olarak 30 megabayt olabilmektedir. Bu yüksek boyutlardan ötürü yapılan çalışmada her bir parçadan 30'ar saniyelik kesitler alınmıştır.

Günümüzde müzik dosyalarının birçoğu sıkıştırılmış dosyalardır. Müzik parçalarının sıkıştırılması farklı müzik formatlarında kullanılmaktadır. Sıkıştırılmış müzik dosyalarının dosya boyutu sıkıştırılma esnasında oldukça düşürülebilmektedir. Ancak wav dosyalarda kayıp olmamasından ve yalnızca sayısallaştırılmış sinyaller içermesinden ötürü, frekansa bağlı öznelikler kolayca elde edilebilmektedir. Wav formatının bu esnekliğinden ötürü veri setinde wav uzantılı müzik dosyaları kullanılmıştır.

### 5.3 Tezin Akış Şeması

Aşağıdaki şekilde tez çalışmasının adım adım aşamaları verilmiştir (Şekil 5.1).



Şekil 5.1 : Tez çalışmasının akış şeması.

Wav formatındaki müzik dosyaları, sınıflandırılmak üzere belirlenmiş 6 türün her biri için 100'er adet olmak üzere webden elde edilerek tüm müzik parçaları “123apps” uygulaması ile kırılarak her bir müzik parçası için 30'ar saniyelik kesitler elde edilmiştir.

Nihaî olarak elde edilen 600 müzik parçası için her birinin python programlama dilinde librosa kütüphanesi kullanılarak; “chroma\_stft”, “rms”, “spectral\_centroid”, “spectral\_bandwidth”, “spectral\_rolloff”, “zero\_crossing\_rate” ve “mfccs” özelliklerinin ortalama değerleri sınıflandırmada kullanılmak üzere csv formatında elde edilmiştir.

Bu veriler standardizasyona tabi tutularak SVM, K-NN ve Lojistik Regresyon ile müzik sınıflandırması yapılmıştır.

İlgili kodlamalar Google'ın ücretsiz erişim sunduğu; Colaboratory, Jupyter Notebook ortamında yapılmıştır. Gerekli kodlamalar için öznitelik çıkarım kısmında python programlama dilinin sunduğu; numpy, pandas, csv, os, pathlib ve librosa kütüphaneleri kullanılmıştır. Elde edilen özniteliklerin işlenerek modellerin oluşturulması ve sınıflandırma işlemlerinin gerçekleştirilmesi için ise sklearn, seaborn, matplotlib, numpy ve pandas kütüphaneleri kullanılmıştır.

Eğitim ve test verileri için çeşitli denemeler sonucu en yüksek performansı verecek şekilde, verilerin %32'si test için kalan %68'i ise eğitim için ayrılmıştır.

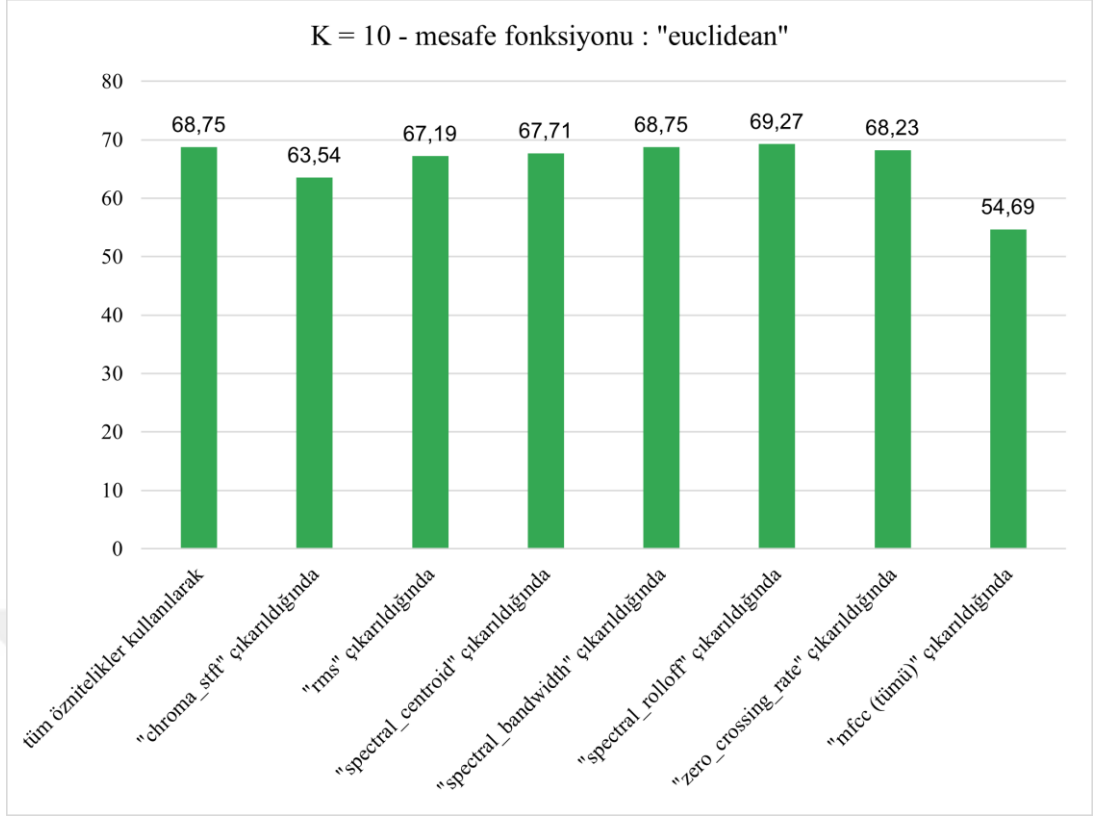
Ses sinyallerinden elde edilen 7 öznitelik her bir algoritma için sınıflandırmaya olan başarımları tek tek gözlemlenmiştir. Elde edilen sonuçların karşılaştırmalı analizleri sunulmuştur.

#### **5.4 K-NN Algoritması ile Müzik Türlerinin Sınıflandırılması**

K-NN algoritması çok yüksek boyutlu olmayan verileri sınıflandırmada iyi performanslar gösterebilmektedir. Bu sebeple veri setine uygun bir algoritma olan K-NN, kullanılacak sınıflandırma algoritmalarından biri olarak belirlenmiştir. Veri setindeki 600 müzik parçası için K-NN algoritması ile “öklid mesafesi”, “manhattan mesafesi” ve “chebyshev mesafesi” fonksiyonları kullanılarak ve başarıma doğrudan etki eden en yakın komşuların belirlendiği K değeri öklid ve manhattan mesafe fonksiyonlarında 10, chebyshev mesafe fonksiyonunda ise 7 olarak belirlenerek müzik türlerinin sınıflandırılması gerçekleştirilmiştir.

“K” değerinin 10 ve 7 seçilmesinin nedeni belirlenen 6 farklı müzik türünün sınıflandırılması için en iyi performansın, K için yapılan birçok deneme sonucunda öklid ve manhattan fonksiyonu için 10, chebyshev fonksiyonu için ise 7 olarak elde edilmiş olmasıdır.

K-NN algoritması ile mesafe fonksiyonu öklid seçildiği durumda ve K değeri 10 olduğu durumda tüm özniteliklerin doğruluk oranlarına etkisi aşağıdaki grafikte verilmiştir (Şekil 5.2).

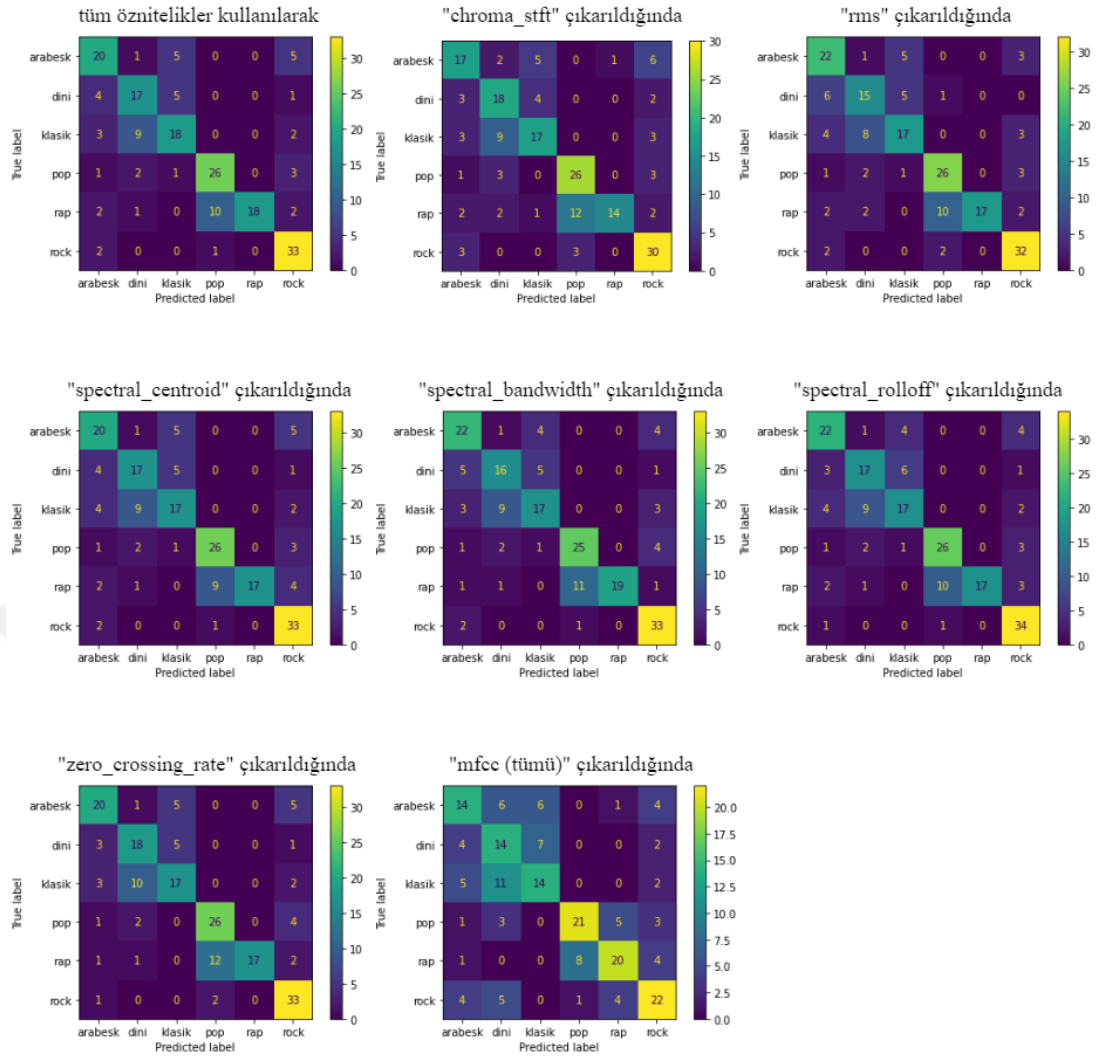


**Şekil 5.2 :** K-NN algoritması ile mesafe fonksiyonu öklid ve K değeri 10 olduğu durumda elde edilen doğruluk oranları.

K-NN algoritması ile öklid mesafe fonksiyonu kullanılarak ve K değeri 10 olarak belirlendiği durumda en yüksek başarıım "spectral\_rolloff" özneliği dahil edilmediği durumda %69.27 olarak görülmüştür. Dolayısıyla bu özneliğin başarıma negatif yönde bir etkisi mevcuttur.

Bunun dışında negatif etki gösteren bir öznelik görülmemiştir. Başarıma en büyük negatif etki ise "mfcc" özneliğinin çıkarıldığı durumda %54.69 doğruluk oranı ile elde edilmiştir. Başarıma en büyük pozitif etkiyi %14.06 gibi bir farkla "mfcc" özneliği sağlamaktadır.

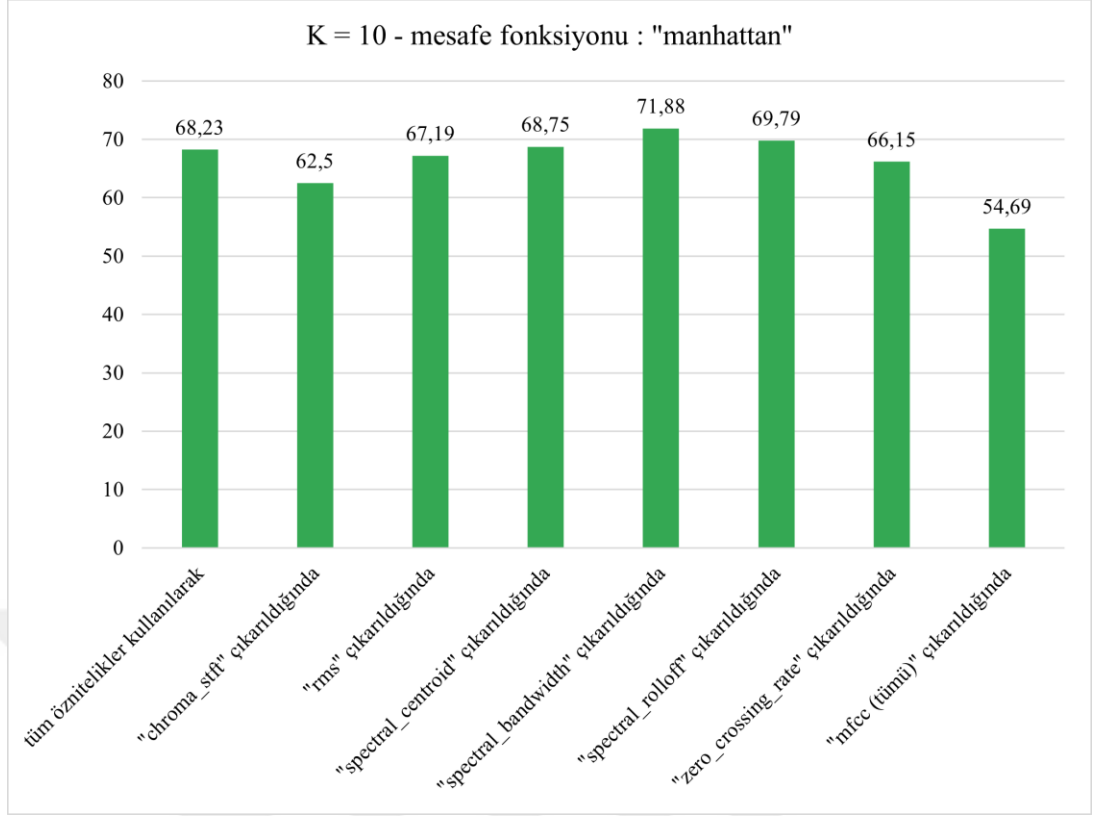
K-NN algoritması ile K=10 alınarak ve öklid mesafe fonksiyonu seçilerek kullanılan özneliklerin tüm durumları için tek tek karmaşıklık matrislerinden elde edilen görüntülerin birleştirilmesi ile elde edilen görsel aşağıda verilmiştir (Şekil 5.3).



**Şekil 5.3 :** K-NN algoritması ile K=10 alınarak ve öklid mesafe fonksiyonu kullanılarak elde edilen karmaşıklık matrisleri.

K-NN algoritması ile K değeri 10 ve mesafe fonksiyonu öklid seçildiği durumda elde edilen karmaşıklık matrislerinden, bu mesafe fonksiyonunun sınıflandırma başarımlarının “dini” ve “klasik” müzik türlerini sınıflandırmada diğer türlere göre daha zayıf kaldığı görülmüştür. “Pop” ve “rock” müzik türlerini sınıflandırmada ise çoğunlukla diğerlerine göre daha yüksek başarımlar elde edilmiştir.

K-NN algoritması ile mesafe fonksiyonu manhattan seçildiği durumda ve K değeri 10 olduğu durumda tüm özelliklerin doğruluk oranlarına etkisi aşağıdaki grafikte verilmiştir (Şekil 5.4).

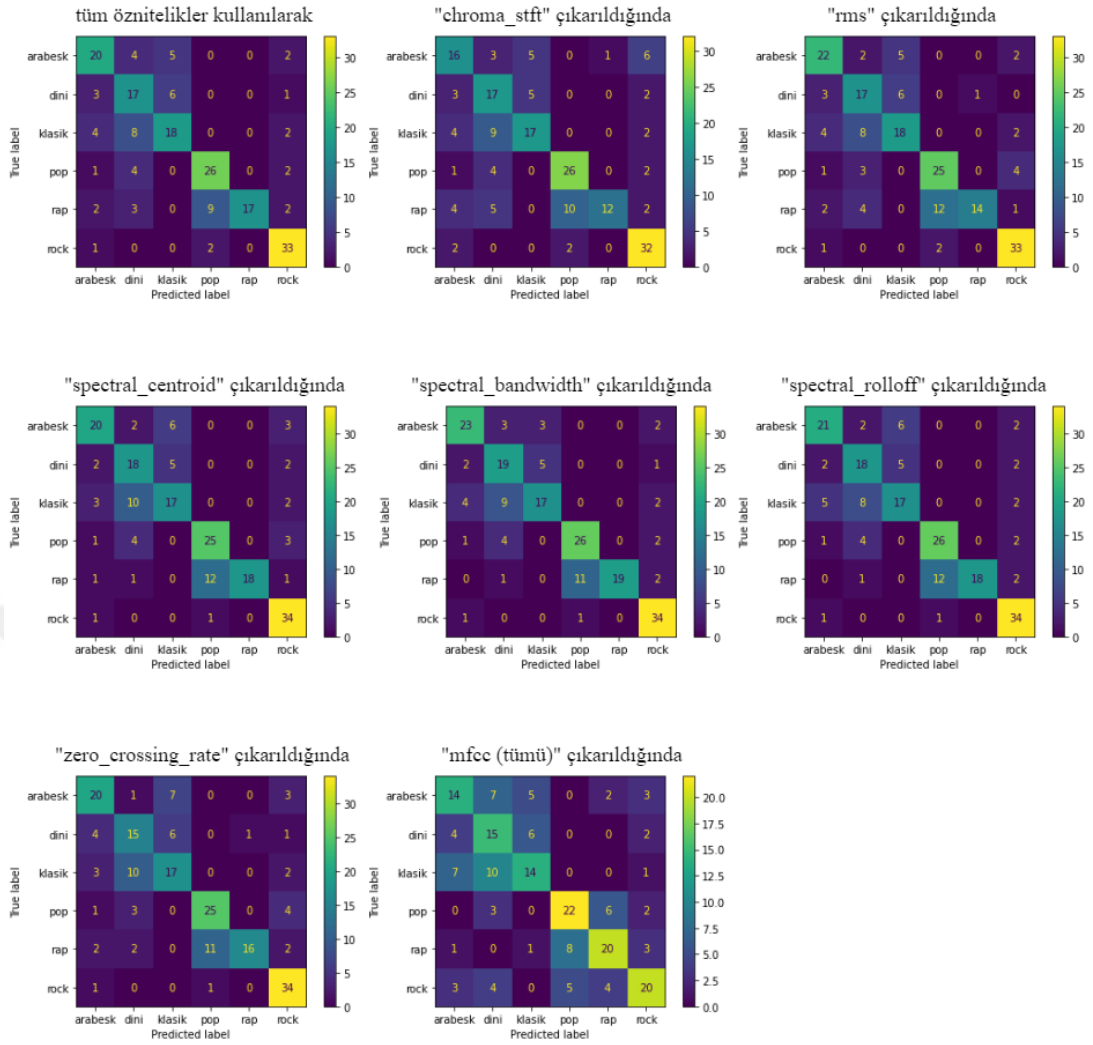


**Şekil 5.4 :** K-NN algoritması ile mesafe fonksiyonu “manhattan” ve K değeri 10 olduğu durumda elde edilen doğruluk oranları.

K-NN algoritması ile “manhattan” mesafe fonksiyonu kullanılarak ve K değeri 10 olarak belirlendiği durumda en yüksek başarıım “spectral\_bandwidth” özneliği dahil edilmediği durumda %71.88 olarak görülmüştür. Dolayısıyla bu özneliğin başarıma negatif yönde bir etkisi mevcuttur.

Bunun dışında “spectral\_centroid” ve “spectral\_rolloff” öznelikleri de başarıma negatif etki göstermiştir. Diğer özneliklerin başarıma pozitif etki sağladığı görülmüştür. Başarıma en büyük negatif etki ise “mfcc” özneliğinin çıkarıldığında tespit edilen en düşük başarıım olan %54.69 doğruluk oranı ile elde edilmiştir. Başarıma en büyük pozitif etkiyi %13.54 gibi bir farkla “mfcc” özneliği sağlamaktadır.

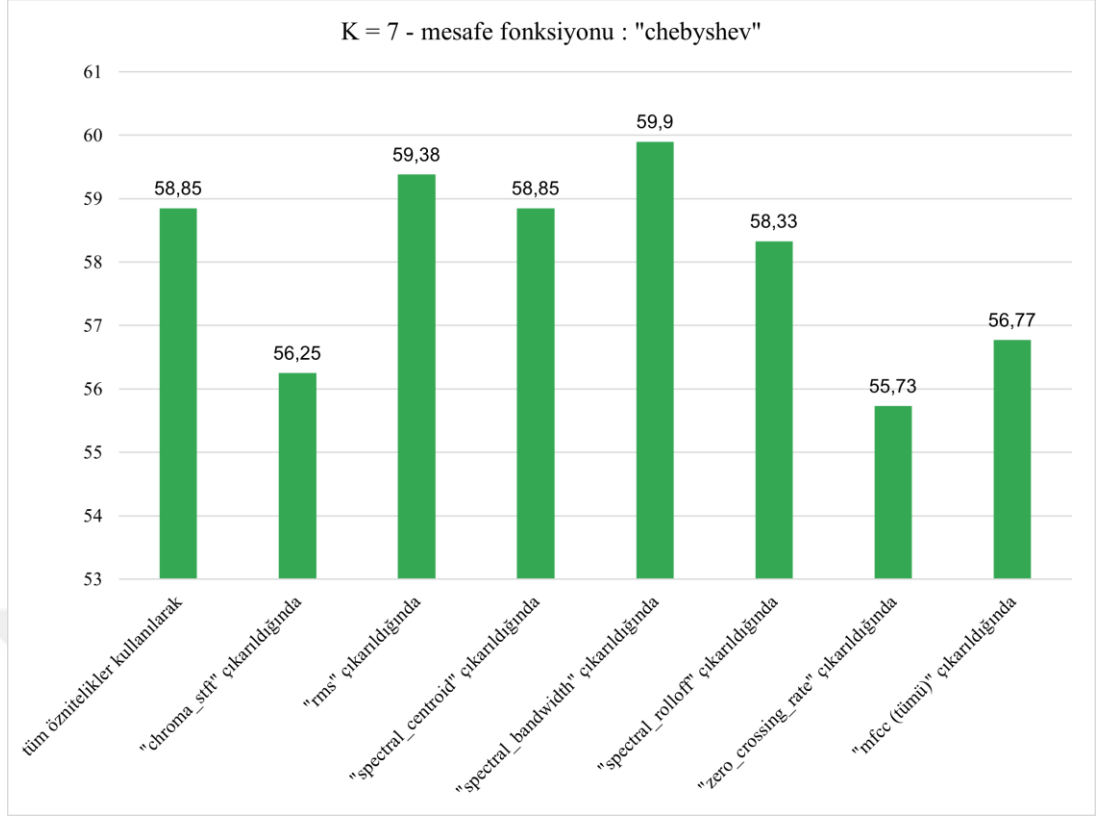
K-NN algoritması ile K=10 alınarak ve manhattan mesafe fonksiyonu seçilerek kullanılan özneliklerin tüm durumları için tek tek karmaşıklık matrislerinden elde edilen görüntülerin birleştirilmesi ile elde edilen görsel aşağıda verilmiştir (Şekil 5.5).



**Şekil 5.5 :** K-NN algoritması ile K=10 alınarak ve “manhattan” mesafe fonksiyonu kullanılarak elde edilen karmaşıklık matrisleri.

K-NN algoritması ile K değeri 10 ve mesafe fonksiyonu “manhattan” seçildiği durumda elde edilen karmaşıklık matrislerinden, bu mesafe fonksiyonunun sınıflandırma başarımlarının “euclidian” mesafe fonksiyonu ile oluşturulan modele benzer şekilde “dini” ve “klasik” müziği sınıflandırmada diğer türlere göre daha zayıf kaldığı görülmüştür. “Pop” ve “rock” müzik türlerini sınıflandırmada ise çoğunlukla diğerlerine göre daha yüksek başarımlar elde edilmiştir.

K-NN algoritması ile mesafe fonksiyonu chebyshev seçildiği durumda ve K değeri 7 olduğu durumda tüm özniteliklerin doğruluk oranlarına etkisi aşağıdaki grafikte verilmiştir (Şekil 5.6).

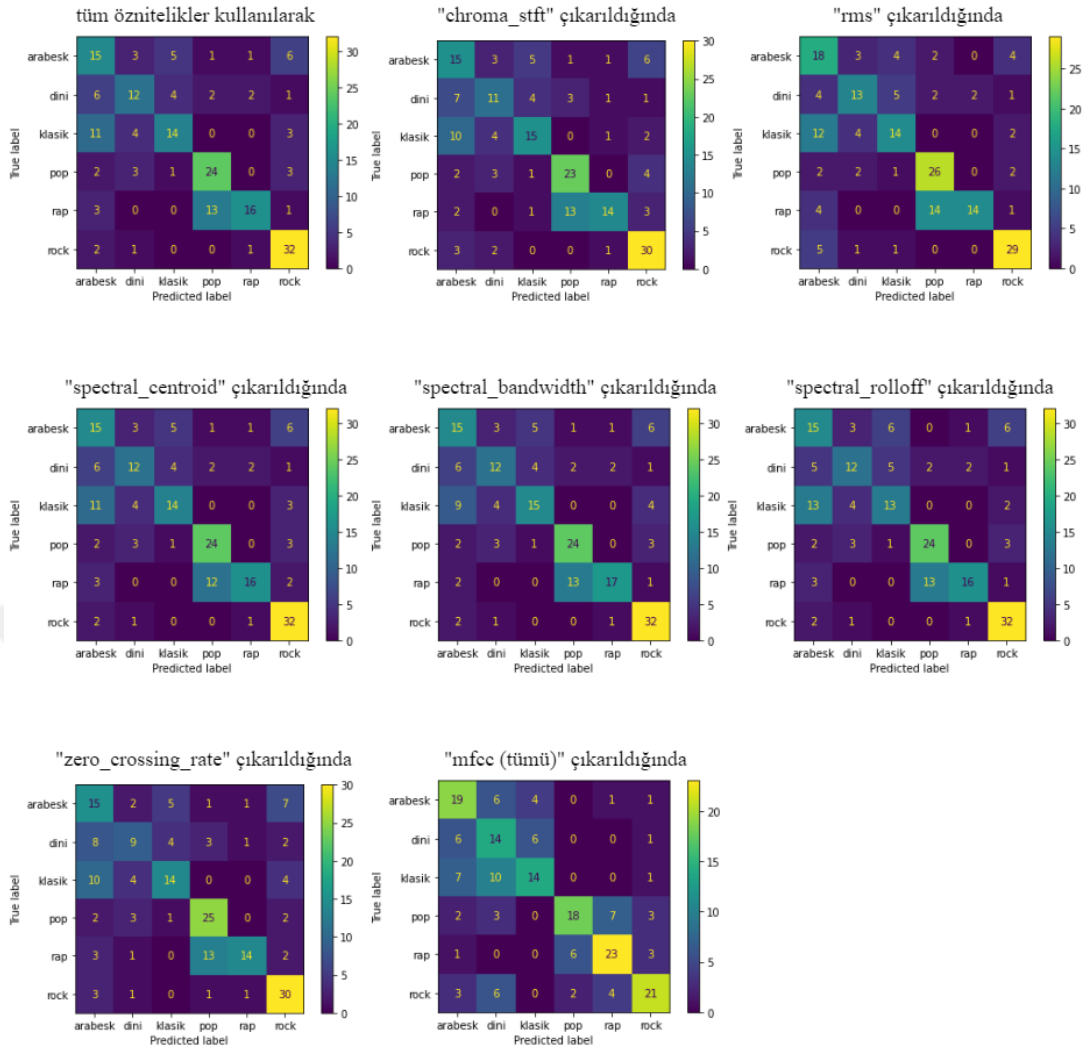


**Şekil 5.6 :** K-NN algoritması ile mesafe fonksiyonu “chebyshev” ve K değeri 7 olduğu durumda elde edilen doğruluk oranları.

K-NN algoritması ile “chebyshev” mesafe fonksiyonu kullanılarak ve K değeri 7 olarak belirlendiği durumda en yüksek başarıım “spectral\_bandwidth” öz niteliği dahil edilmediği durumda %59.9 olarak görülmüştür. Dolayısıyla bu öz niteliğin başarıma etkisi negatiftir.

Bunun dışında “rms” öz niteliği de başarıma negatif etki göstermiştir. Başarıma nötr etki gösteren “spectral\_centroid” öz niteliği dahil edilmediği durumda %58.85 doğruluk oranı görülmüştür. Diğer öz niteliklerin başarıma pozitif etki sağladığı görülmüştür. Başarıma en büyük negatif etki ise diğer mesafe fonksiyonlarının aksine “zero\_crossing\_rate” öz niteliğinin çıkarıldığında durumda, en düşük başarıım olan %55.73 doğruluk oranı ile elde edilmiştir. Başarıma en büyük pozitif etkiyi “zero\_crossing\_rate” öz niteliği sağlamaktadır.

K-NN algoritması ile K=7 alınarak ve chebyshev mesafe fonksiyonu seçilerek kullanılan öz niteliklerin tüm durumları için tek tek karmaşıklık matrislerinden elde edilen görüntülerin birleştirilmesi ile elde edilen görsel aşağıda verilmiştir (Şekil 5.7).



**Şekil 5.7 :** K-NN algoritması ile K=7 alınarak ve “chebyshev” mesafe fonksiyonu kullanılarak elde edilen karmaşıklık matrisleri.

K-NN algoritması ile K değeri 7 ve mesafe fonksiyonu “chebyshev” seçildiği durumda elde edilen karmaşıklık matrislerinden, bu mesafe fonksiyonunun sınıflandırma başarımlarının genel olarak en yüksek “rock” müzik türünde gözlemlendiği tespit edilmiştir. Yalnızca “mfcc” öznitelikleri çıkarıldığı durumda rap müzik türünün sınıflandırma başarımları en daha yüksek olarak görülmüştür. “Arabesk”, “dini” ve “klasik” müzik türlerinde ise diğer türlere göre sınıflandırma başarımları zayıf kalmıştır.

K-NN algoritması tüm mesafe fonksiyonlarında genel olarak en fazla “dini” ve “klasik” müzik türlerini birbirinden ayırt etmede zorlanmıştır. Müzik türlerinin duygusal temalarının benzer olması bu türlerin birbirinden ayrıştırılmasını önemli ölçüde zorlaştırmaktadır.

K-NN ile öklid mesafe fonksiyonu ile en yüksek başarıyı veren öznitelik seçimi ile sınıflandırılan türlerin kesinlik, duyarlılık ve F1-skor değerleri aşağıda verilmiştir (Çizelge 5.1).

**Çizelge 5.1 :** K-NN ile öklid mesafe fonksiyonu kullanılarak elde edilen en yüksek başarıyı veren öznitelik seçimi ile sınıflandırılan türlerin başarılarını gösteren metrikler.

Metrikler	Precision	Recall	F1-Score
arabesk	0.67	0.71	0.69
dini	0.57	0.63	0.60
klasik	0.61	0.53	0.57
pop	0.70	0.79	0.74
rap	1.00	0.52	0.68
rock	0.72	0.94	0.82
Ağırlıklı Ortalama	0.72	0.69	0.69

K-NN algoritması ile öklid mesafe fonksiyonu kullanıldığı durumda en yüksek precision (kesinlik) değeri “rap” müzik türünde “1.00” ile elde edilmiştir. Pozitif tahminlerin tümünün gerçekte de pozitif olduğu görülmektedir. En yüksek recall (duyarlılık) değeri ise “rock” müzik türünde elde edilmiştir. Pozitif tahmin edilmesi beklenen durumların yaklaşık %94’ü doğru tahmin edilmiştir. En yüksek F1-Skor ise yine “rock” müzik türünde elde edilmiştir. Ağırlıklı ortalaması en yüksek olan metrik “kesinlik” olarak görülmüştür.

K-NN ile manhattan mesafe fonksiyonu ile en yüksek başarıyı veren öznitelik seçimi ile sınıflandırılan türlerin kesinlik, duyarlılık ve F1-skor değerleri aşağıda verilmiştir (Çizelge 5.2).

**Çizelge 5.2 :** K-NN ile manhattan mesafe fonksiyonu kullanılarak elde edilen en yüksek başarıyı veren öznitelik seçimi ile sınıflandırılan türlerin başarılarını gösteren metrikler.

Metrikler	Precision	Recall	F1-Score
arabesk	0.74	0.74	0.74
dini	0.53	0.70	0.60
klasik	0.68	0.53	0.60
pop	0.68	0.79	0.73
rap	1.00	0.58	0.73
rock	0.79	0.94	0.86
Ağırlıklı Ortalama	0.75	0.72	0.72

K-NN algoritması ile manhattan mesafe fonksiyonu kullanıldığı durumda elde edilen en yüksek değere sahip olan metrikler, öklid mesafe fonksiyonu kullanıldığı

durumdaki ile aynıdır. Yalnızca az miktarda bir farklılık söz konusudur. Ağırlıklı ortalaması en yüksek olan metriğin yine “kesinlik” olduğu görülmüştür.

K-NN ile chebyshev mesafe fonksiyonu ile en yüksek başarıyı veren öznelik seçimi ile sınıflandırılan türlerin kesinlik, duyarlılık ve F1-skor değerleri aşağıda verilmiştir (Çizelge 5.3).

**Çizelge 5.3 :** K-NN ile chebyshev mesafe fonksiyonu kullanılarak elde edilen en yüksek başarıyı veren öznelik seçimi ile sınıflandırılan türlerin başarımlarını gösteren metrikler.

Metrikler	Precision	Recall	F1-Score
arabesk	0.42	0.48	0.45
dini	0.55	0.44	0.48
klasik	0.60	0.47	0.53
pop	0.60	0.73	0.66
rap	0.81	0.52	0.63
rock	0.68	0.89	0.77
Ağırlıklı Ortalama	0.61	0.60	0.59

K-NN algoritması ile chebyshev mesafe fonksiyonu kullanıldığı durumda en yüksek precision (kesinlik) değeri “0.81” ile rap müzik türünde elde edilmiştir. Pozitif tahminlerin %81’inin pozitif olduğu görülmektedir. En yüksek recall (duyarlılık) değeri ise 0.89 ile “rock” müzik türünde elde edilmiştir. Pozitif tahmin edilmesi beklenen durumların yaklaşık %89’u doğru tahmin edilmiştir. En yüksek F1-Skor ise yine “rock” müzik türünde elde edilmiştir. Ağırlıklı ortalaması en yüksek olan metrik 0.61 ile “kesinlik” olarak görülmüştür.

K-NN ile arabesk müzik türü için tüm mesafe fonksiyonlarından elde edilen en yüksek kesinlik, duyarlılık ve F1-skor değerleri aşağıda verilmiştir (Çizelge 5.4).

**Çizelge 5.4 :** K-NN ile arabesk müzik türü için mesafe fonksiyonlarına ait metriklerin en yüksek başarımları.

Tür : Arabesk	Precision	Recall	F1-Score
Öklid	0.67	0.71	0.69
Manhattan	0.74	0.74	0.74
Chebyshev	0.42	0.48	0.45

K-NN algoritması ile tüm mesafe fonksiyonları kullanılarak “Arabesk” müzik türünde elde edilen en yüksek değer “kesinlik” için 0.74, “duyarlılık” için 0.74, “F1 Skor” için ise yine 0.74 olarak görülmüştür.

K-NN ile dini müzik türü için tüm mesafe fonksiyonlarından elde edilen en yüksek kesinlik, duyarlılık ve F1-skor değerleri aşağıda verilmiştir (Çizelge 5.5).

**Çizelge 5.5 :** K-NN ile dini müzik türü için mesafe fonksiyonlarına ait metriklerin en yüksek başarımları.

Tür : Dini	Precision	Recall	F1-Score
Öklid	0.57	0.63	0.60
Manhattan	0.53	0.70	0.60
Chebyshev	0.55	0.44	0.48

K-NN algoritması ile tüm mesafe fonksiyonları kullanılarak “Dini” müzik türünde elde edilen en yüksek değer “kesinlik” için 0.57, “duyarlılık” için 0.70, “F1 Skor” için ise yine 0.60 olarak görülmüştür.

K-NN ile klasik müzik türü için tüm mesafe fonksiyonlarından elde edilen en yüksek kesinlik, duyarlılık ve F1-skor değerleri aşağıda verilmiştir (Çizelge 5.6).

**Çizelge 5.6 :** K-NN ile klasik müzik türü için mesafe fonksiyonlarına ait metriklerin en yüksek başarımları.

Tür : Klasik	Precision	Recall	F1-Score
Öklid	0.61	0.53	0.57
Manhattan	0.68	0.53	0.60
Chebyshev	0.60	0.47	0.53

K-NN algoritması ile tüm mesafe fonksiyonları kullanılarak “Klasik” müzik türünde elde edilen en yüksek değer “kesinlik” için 0.68, “duyarlılık” için 0.56, “F1 Skor” için ise yine 0.60 olarak görülmüştür.

K-NN ile pop müzik türü için tüm mesafe fonksiyonlarından elde edilen en yüksek kesinlik, duyarlılık ve F1-skor değerleri aşağıda verilmiştir (Çizelge 5.7).

**Çizelge 5.7 :** K-NN ile pop müzik türü için mesafe fonksiyonlarına ait metriklerin en yüksek başarımları.

Tür : Pop	Precision	Recall	F1-Score
Öklid	0.70	0.79	0.74
Manhattan	0.68	0.79	0.73
Chebyshev	0.60	0.73	0.66

K-NN algoritması ile tüm mesafe fonksiyonları kullanılarak “Pop” müzik türünde elde edilen en yüksek değer “kesinlik” için 0.70, “duyarlılık” için 0.79, “F1 Skor” için ise yine 0.74 olarak görülmüştür.

K-NN ile rap müzik türü için tüm mesafe fonksiyonlarından elde edilen en yüksek kesinlik, duyarlılık ve F1-skor değerleri aşağıda verilmiştir (Çizelge 5.8).

**Çizelge 5.8 :** K-NN ile rap müzik türü için mesafe fonksiyonlarına ait metriklerin en yüksek başarımları.

Tür : Rap	Precision	Recall	F1-Score
Öklid	1.00	0.52	0.68
Manhattan	1.00	0.58	0.73
Chebyshev	0.81	0.52	0.63

K-NN algoritması ile tüm mesafe fonksiyonları kullanılarak “Rap” müzik türünde elde edilen en yüksek değer “kesinlik” için 1.00, “duyarlılık” için 0.58, “F1 Skor” için ise yine 0.73 olarak görülmüştür.

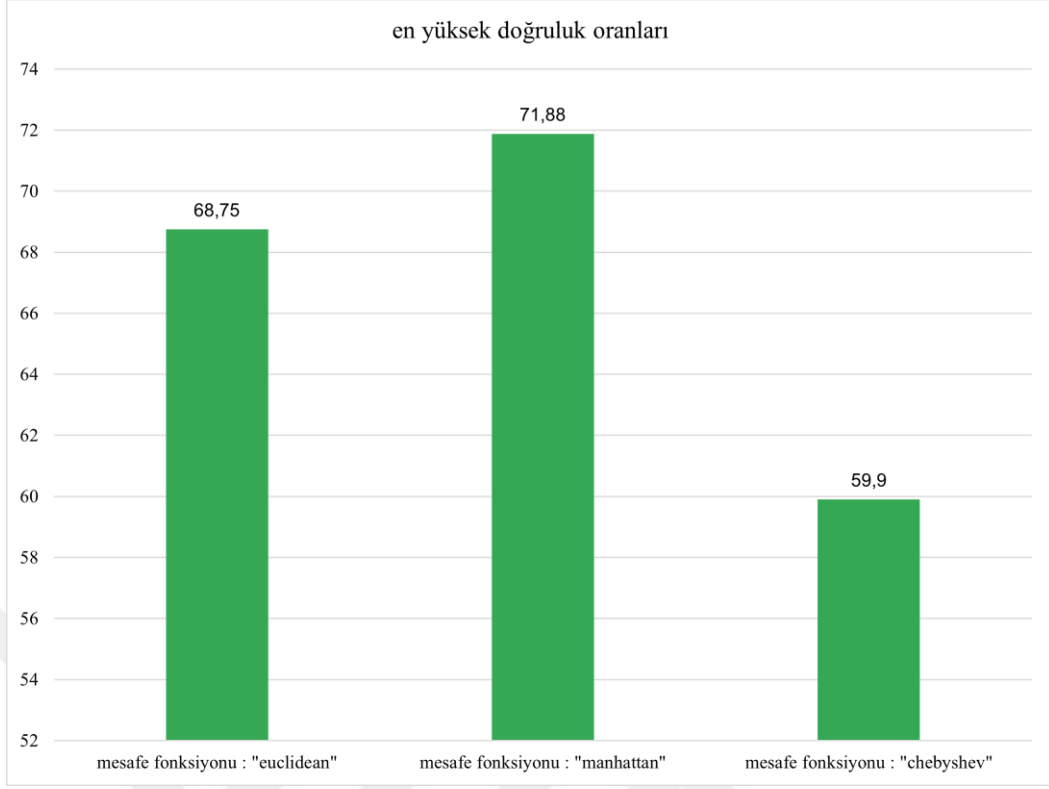
K-NN ile rock müzik türü için tüm mesafe fonksiyonlarından elde edilen en yüksek kesinlik, duyarlılık ve F1-skor değerleri aşağıda verilmiştir (Çizelge 5.9).

**Çizelge 5.9 :** K-NN ile rock müzik türü için mesafe fonksiyonlarına ait metriklerin en yüksek başarımları.

Tür : Rock	Precision	Recall	F1-Score
Öklid	0.72	0.94	0.82
Manhattan	0.79	0.94	0.86
Chebyshev	0.68	0.89	0.77

K-NN algoritması ile tüm mesafe fonksiyonları kullanılarak “Rock” müzik türünde elde edilen en yüksek değer “kesinlik” için 0.79, “duyarlılık” için 0.94, “F1 Skor” için ise yine 0.86 olarak görülmüştür.

K-NN algoritması ile mesafe fonksiyonu seçimine göre elde edilen en yüksek doğruluk oranları aşağıdaki grafikte verilmiştir (Şekil 5.8).



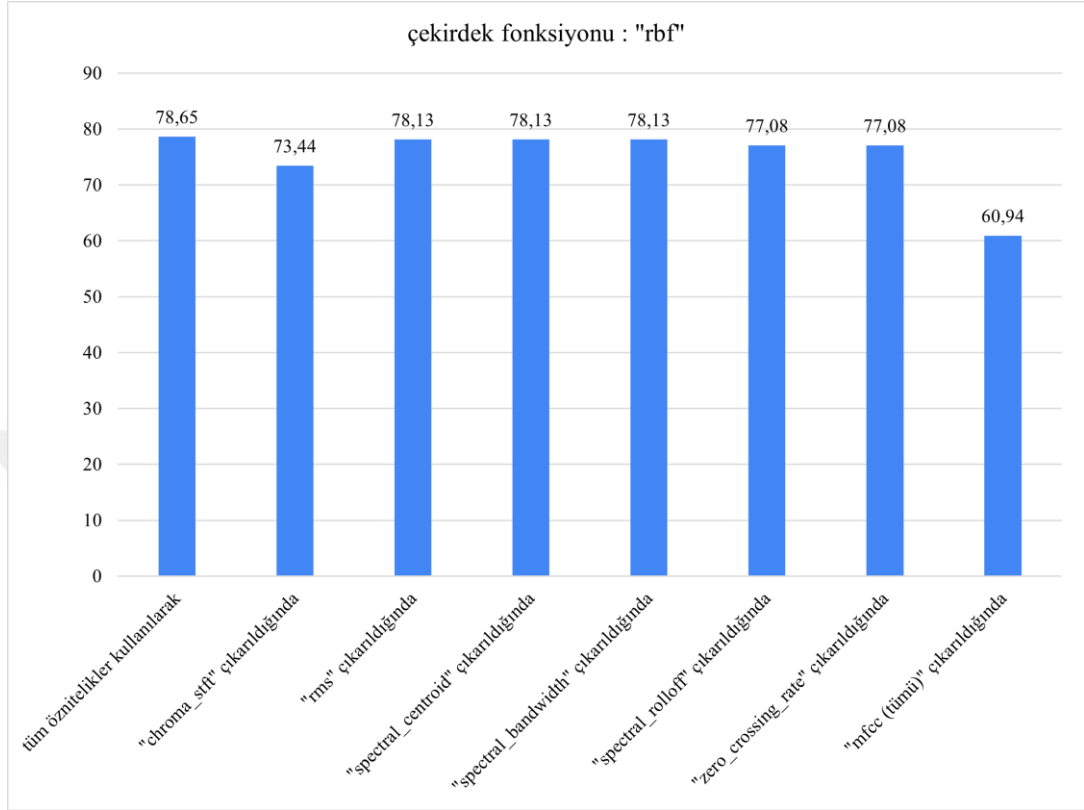
**Şekil 5.8 :** K-NN algoritması ile seçilen mesafe fonksiyonuna göre elde edilen en yüksek doğruluk oranları.

K-NN ile yakalanan en yüksek başarımların "manhattan" mesafe fonksiyonu ile "spectral\_bandwidth" özneliği çıkarıldığında elde edilmiştir. "Öklid" mesafe fonksiyonu kullanıldığı durumda en yüksek başarımların tüm özneliklerin kullanılarak %68.75 olduğu görülmüştür. "Chebyshev" mesafe fonksiyonunun kullanıldığı durumda ise en yüksek başarımların "zero\_crossing\_rate" özneliği çıkarıldığı durumda %59.9 ile elde edilmiştir.

### 5.5 SVM Algoritması ile Müzik Türlerinin Sınıflandırılması

Müzik türü sınıflandırmasında başarılı performanslar gösteren SVM algoritması ile "radyal tabanlı çekirdek", "polinomal çekirdek" ve "doğrusal çekirdek" olmak üzere 3 farklı çekirdek fonksiyonu kullanılarak modeller oluşturulup sınıflandırma işlemleri gerçekleştirilmiştir. Tüm özneliklerin tek tek başarımlarına olan etkisi tespit edilmek üzere sırasıyla özneliklerden her biri hariç tutularak sınıflandırmadaki doğruluk oranları gözlemlenmiştir.

Aşağıdaki şekilde, SVM algoritması ile radyal tabanlı çekirdek fonksiyonu kullanıldığı durumda tüm özniteliklerin doğruluk oranlarına etkisi gösterilmiştir (Şekil 5.9).

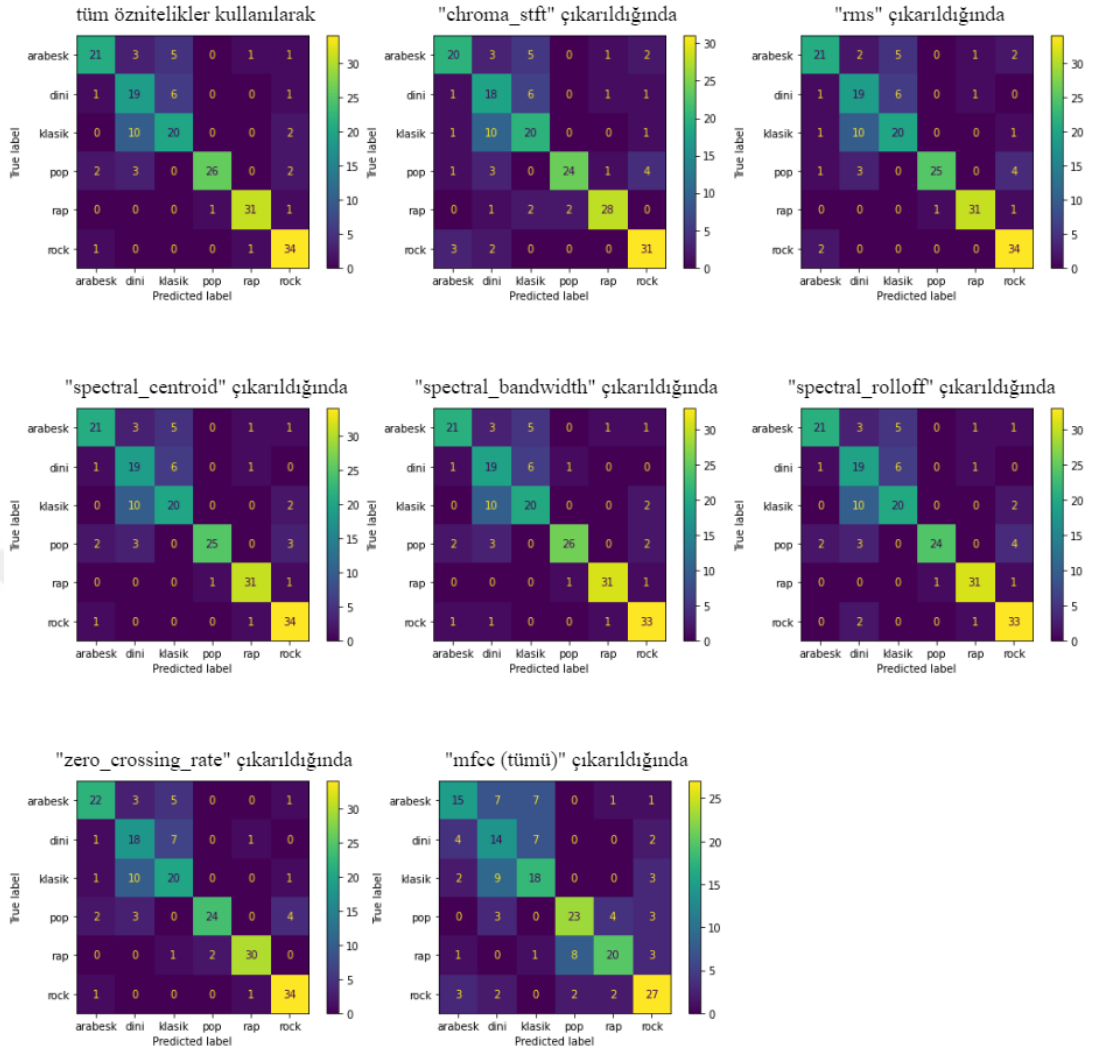


**Şekil 5.9 :** SVM algoritması ile radyal tabanlı çekirdek fonksiyonuna ait doğruluk oranları.

SVM algoritması ile radyal tabanlı çekirdek fonksiyonu kullanıldığı durumda başarıma en büyük etkisi olan özneliğin "mfcc" olduğu görülmüştür. "Mfcc" başarıma pozitif yönde etki etmektedir. Dolayısıyla en düşük doğruluk oranı "mfcc" dahil edilmediği durumda görülmüştür. Bununla birlikte "chroma\_stft" özneliğinin de başarıma pozitif yönde %5.21 gibi ciddi bir etkisi olduğu tespit edilmiştir.

Kalan diğer özneliklerin çok büyük bir etkisi olmadığı en fazla %1-2 civarında bir etkileri olduğu gözlemlenmiştir. En yüksek doğruluk oranı ise tüm özneliklerin kullanıldığı durumda %78.65 olarak elde edilmiştir. Radyal tabanlı çekirdek fonksiyonu ile oluşturulan modelde tüm özneliklerin başarıma pozitif etki ettiği tespit edilmiştir.

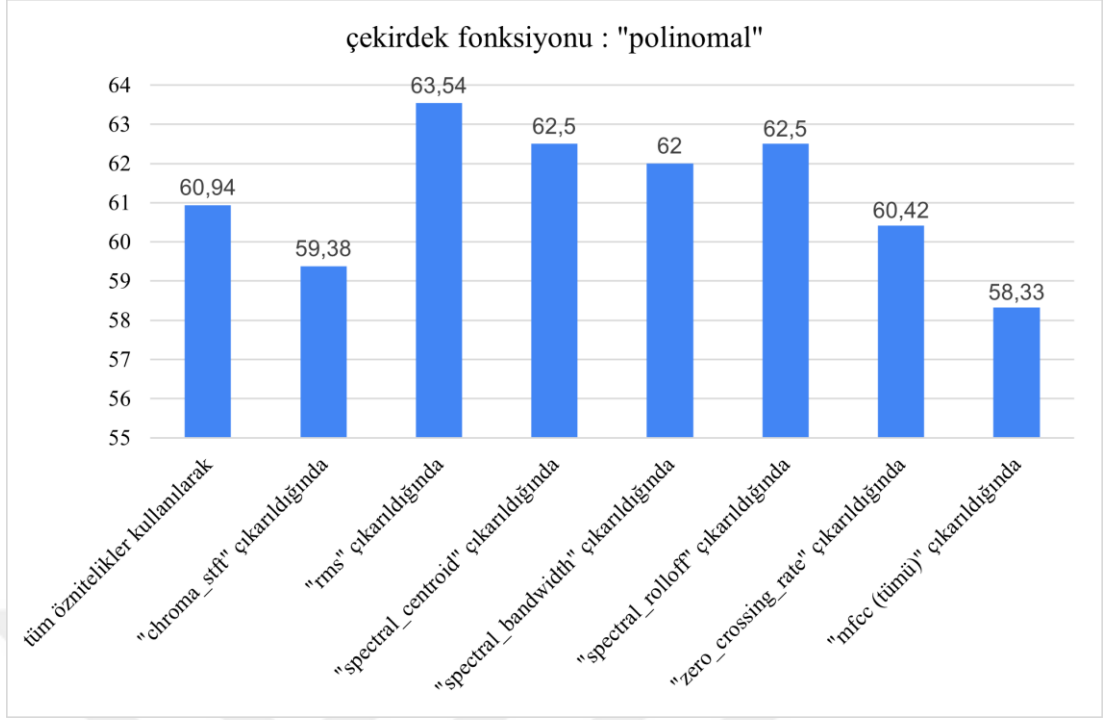
SVM algoritması ile radyal tabanlı çekirdek fonksiyonu kullanılarak özneliklerin tüm durumları için tek tek karmaşıklık matrislerinden elde edilen görüntülerin birleştirilmesi ile elde edilen görsel aşağıda verilmiştir (Şekil 5.10).



**Şekil 5.10** : SVM algoritması ile radyal tabanlı çekirdek fonksiyonu kullanılarak elde edilen karmaşıklık matrisleri.

SVM algoritması ile radyal tabanlı çekirdek fonksiyonu kullanıldığı durumda elde edilen karmaşıklık matrislerinden, ayrı ayrı kombinasyonel olarak kullanılan tüm özniteliklerin bir genellemesini yaptığımız durumda; “arabesk”, “dini” ve “klasik” müziği sınıflandırmada nispeten başarısız olduğu görülmüştür. “Rap” ve “rock” müzik türlerini sınıflandırmada ise çoğunlukla diğerlerine göre daha yüksek başarımlar elde edilmiştir.

Aşağıdaki şekilde, SVM algoritması ile polinomal çekirdek fonksiyonu kullanıldığı durumda tüm özniteliklerin doğruluk oranlarına etkisi gösterilmiştir (Şekil 5.11).

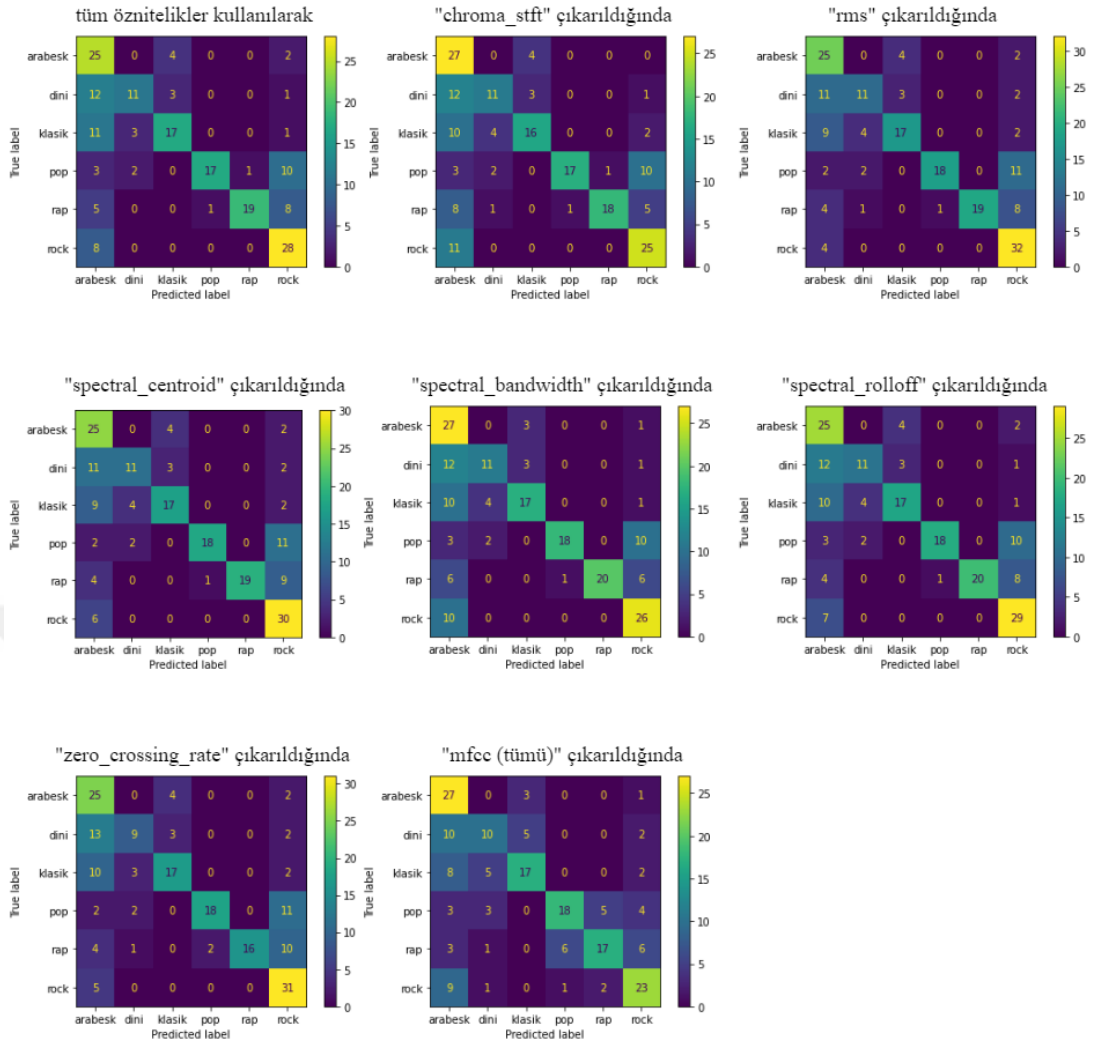


**Şekil 5.11** : SVM algoritması ile polinomal çekirdek fonksiyonuna ait doğruluk oranları.

SVM algoritması ile polinomal çekirdek fonksiyonu kullanılarak elde edilen sonuçlarda en yüksek başarımın, "rms" özneliğinin çıkarıldığı durumda %63.56 doğruluk oranı ile yakalandığı tespit edilmiştir. Dolayısıyla başarıma negatif yönde en fazla etki eden öznelik "rms"dir. Başarımı pozitif yönde en fazla etkileyen öznelikler ise sırasıyla "mfcc" ve buna yakın bir oranda etki gösteren "chroma\_stft" olarak gözlemlenmiştir. Bu 2 öznelik de ayrı ayrı dahil edilmediği durumda doğruluk oranı %60'ın altına kadar düşmüştür.

Ayrıca izgesel (spectral) özneliklerin dahil edilmediği durumlarda başarımın arttığı, dolayısıyla polinomal çekirdek fonksiyonu kullanıldığında izgesel özneliklerin tümünün başarıma negatif yönde etki ettiği görülmüştür.

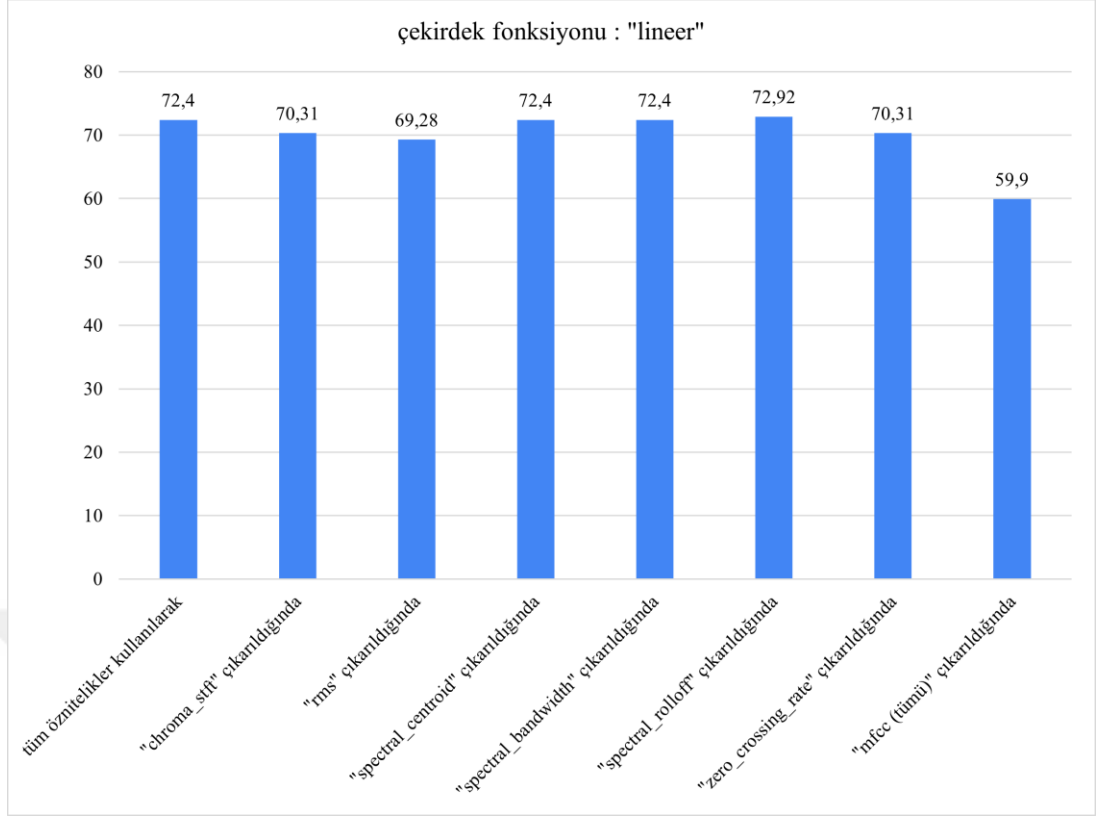
SVM algoritması ile polinomal çekirdek fonksiyonu kullanılarak özneliklerin tüm durumları için tek tek karmaşıklık matrislerinden elde edilen görüntülerin birleştirilmesi ile elde edilen görsel aşağıda verilmiştir (Şekil 5.12).



**Şekil 5.12** : SVM algoritması ile polinomal çekirdek fonksiyonu kullanılarak elde edilen karmaşıklık matrisleri.

SVM algoritması ile polinomal çekirdek fonksiyonu kullanıldığında elde edilen karmaşıklık matrislerinden, ayrı ayrı kombinasyonel olarak kullanılan tüm özneliklerin bir genellemesini yaptığımız durumda; “dini” müzik türünü doğru tahmin etme konusunda daha zayıf kaldığı, “arabesk” ve “rock” müzik türlerinde ise nispeten daha fazla doğru tahminler ürettiği görülmektedir. Diğer türlerin tahminlerinin doğruluk durumlarının birbirine oldukça yakın olduğu görülmüştür.

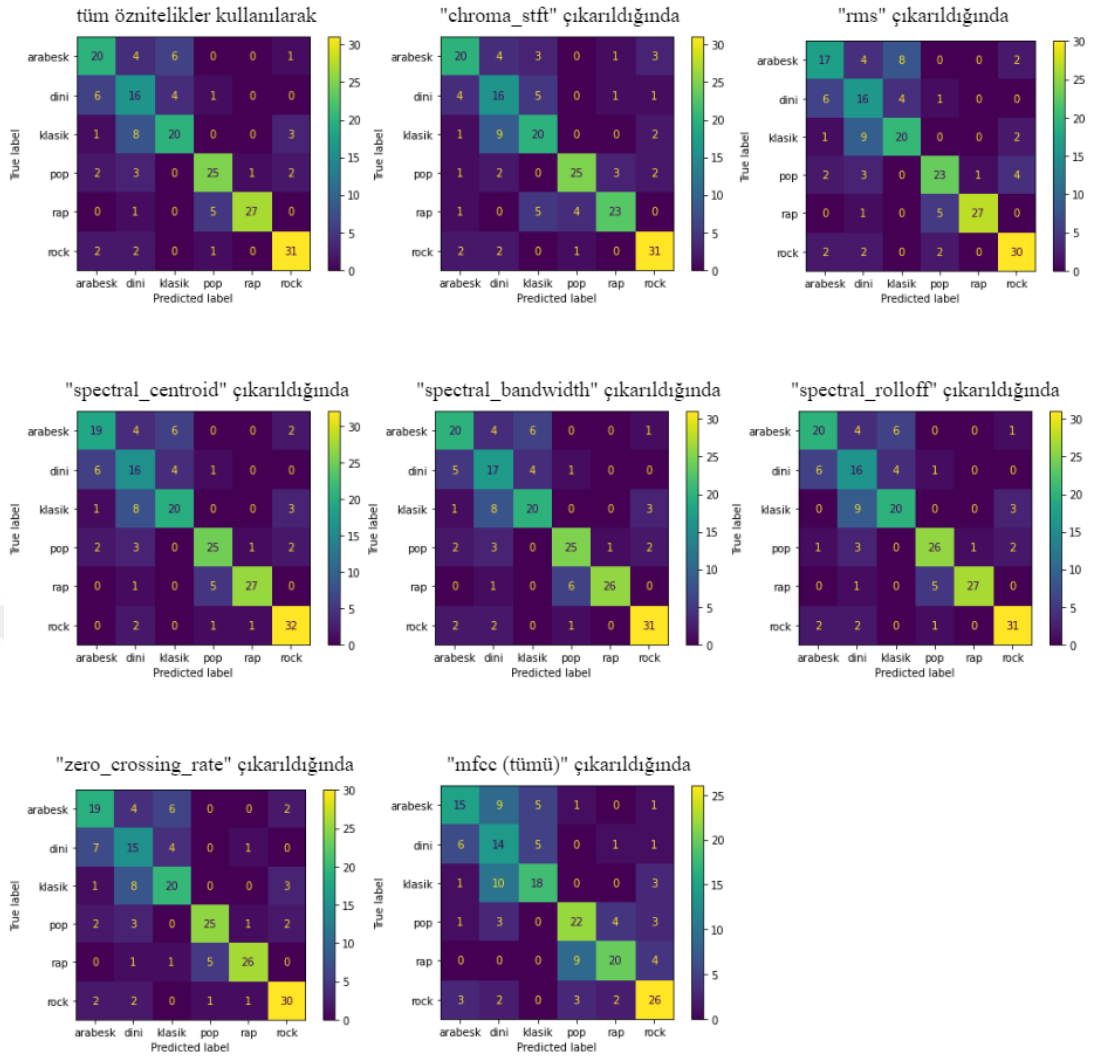
Aşağıdaki şekilde, SVM algoritması ile lineer çekirdek fonksiyonu kullanıldığı durumda tüm özneliklerin doğruluk oranlarına etkisi gösterilmiştir (Şekil 5.13).



**Şekil 5.13 :** SVM algoritması ile lineer çekirdek fonksiyonuna ait doğruluk oranları.

SVM algoritması ile lineer çekirdek fonksiyonu kullanılarak uygulanan modelde "spectral\_rolloff" özneliğinin çıkarıldığı durumda en yüksek başarımın %72.92 doğruluk oranı ile elde edildiği görülmüştür. Dolayısıyla bu öznelik başarıma en büyük negatif etkiye sahiptir. Bu çekirdek fonksiyonu ile yalnızca "spectral\_rolloff" özneliği başarıma negatif yönde etki göstermiştir. Başarıma pozitif en büyük etkiyi ise "mfcc" özneliğinin sağladığı görülmüştür. "Mfcc" dahil edilmediği durumda %59.9 doğruluk oranı ile en düşük başarım gözlemlenmiştir. Lineer çekirdek fonksiyonu uygulandığında yalnızca "mfcc"nin başarıma etkisi %12,5 gibi yüksek bir değer olarak görülmüştür.

SVM algoritması ile lineer çekirdek fonksiyonu kullanılarak özneliklerin tüm durumları için tek tek karmaşıklık matrislerinden elde edilen görüntülerin birleştirilmesi ile elde edilen görsel aşağıda verilmiştir (Şekil 5.14).



**Şekil 5.14 :** SVM algoritması ile linear çekirdek fonksiyonu kullanılarak elde edilen karmaşıklık matrisleri.

SVM algoritması ile linear çekirdek fonksiyonu kullanıldığında elde edilen karmaşıklık matrislerinden, ayrı ayrı kombinasyonel olarak kullanılan tüm özniteliklerin bir genellemesini yaptığımız durumda bu çekirdek fonksiyonunun en yüksek doğru tahmin oranlarını; “pop”, “rap”, “rock” müzik türlerinde yakalandığı gözlemlenmiştir. “Dini” müzik türünde ise en düşük doğru tahmin oranları görülmüştür.

SVM algoritması ile her üç çekirdek fonksiyonu kullanılarak genel sonuçlara baktığımızda bu fonksiyonların her birinin başarımlarının birbirinden farklı olduğu ancak türler arasında en fazla “dini” müzik türünü ayırt etmede zorlandığı görülmüştür. En doğru şekilde tahmin edilen türler ise çoğunlukla “rap” ve “rock” müzik türleri olmuştur.

SVM algoritması ile radyal tabanlı çekirdek fonksiyonu kullanılarak en yüksek başarıyı veren öznitelik seçimi ile sınıflandırılan türlerin kesinlik, duyarlılık ve F1-skor değerleri aşağıda verilmiştir (Çizelge 5.10).

**Çizelge 5.10 :** SVM ile radyal tabanlı çekirdek fonksiyonu kullanılarak elde edilen en yüksek başarıyı veren öznitelik seçimi ile sınıflandırılan türlerin başarımlarını gösteren metrikler.

Metrikler	Precision	Recall	F1-Score
arabesk	0.84	0.68	0.75
dini	0.54	0.70	0.61
klasik	0.65	0.62	0.63
pop	0.96	0.79	0.87
rap	0.94	0.94	0.94
rock	0.83	0.94	0.88
Ağırlıklı Ortalama	0.80	0.79	0.79

SVM algoritması ile radyal tabanlı çekirdek fonksiyonu kullanıldığı durumda en yüksek precision (kesinlik) değeri “pop” müzik türünde 0.96 ile elde edilmiştir. Pozitif tahminlerin yaklaşık olarak %96’sı gerçekten pozitiftir. En yüksek recall (duyarlılık) değeri ise “rap” ve “rock” müzik türünde 0.94 ile elde edilmiştir. Pozitif tahmin edilmesi beklenen durumların yaklaşık %94’ü doğru tahmin edilmiştir. En yüksek F1-Skor ise 0.94 ile “rap” müzik türünde elde edilmiştir. Ağırlıklı ortalaması en yüksek olan metrik “kesinlik” olarak görülmüştür.

SVM algoritması ile polinomal çekirdek fonksiyonu kullanılarak en yüksek başarıyı veren öznitelik seçimi ile sınıflandırılan türlerin kesinlik, duyarlılık ve F1-skor değerleri aşağıda verilmiştir (Çizelge 5.11).

**Çizelge 5.11 :** SVM ile polinomal çekirdek fonksiyonu kullanılarak elde edilen en yüksek başarıyı veren öznitelik seçimi ile sınıflandırılan türlerin başarımlarını gösteren metrikler.

Metrikler	Precision	Recall	F1-Score
arabesk	0.45	0.81	0.58
dini	0.61	0.41	0.49
klasik	0.71	0.53	0.61
pop	0.95	0.55	0.69
rap	1.00	0.58	0.73
rock	0.83	0.94	0.88
Ağırlıklı Ortalama	0.72	0.64	0.64

SVM algoritması ile polinomal çekirdek fonksiyonu kullanıldığı durumda en yüksek precision (kesinlik) değeri “rap” müzik türünde 1.00 ile elde edilmiştir. Pozitif

tahminlerin tümü gerçekten pozitiftir. En yüksek recall (duyarlılık) değeri ise “rock” müzik türünde 0.94 ile elde edilmiştir. Pozitif tahmin edilmesi beklenen durumların yaklaşık %94’ü doğru tahmin edilmiştir. En yüksek F1-Skor ise 0.88 ile “rock” müzik türünde elde edilmiştir. Ağırlıklı ortalaması en yüksek olan metrik “kesinlik” olarak görülmüştür.

SVM algoritması ile lineer çekirdek fonksiyonu kullanılarak en yüksek başarıyı veren öznelik seçimi ile sınıflandırılan türlerin kesinlik, duyarlılık ve F1-skor değerleri aşağıda verilmiştir (Çizelge 5.12).

**Çizelge 5.12** : SVM ile lineer çekirdek fonksiyonu kullanılarak elde edilen en yüksek başarıyı veren öznelik seçimi ile sınıflandırılan türlerin başarımlarını gösteren metrikler.

Metrikler	Precision	Recall	F1-Score
arabesk	0.69	0.65	0.67
dini	0.46	0.59	0.65
klasik	0.67	0.62	0.65
pop	0.79	0.79	0.79
rap	0.96	0.82	0.89
rock	0.84	0.86	0.85
Ağırlıklı Ortalama	0.74	0.73	0.73

SVM algoritması ile lineer çekirdek fonksiyonu kullanıldığı durumda en yüksek precision (kesinlik) değeri “rap” müzik türünde 0.96 ile elde edilmiştir. Pozitif tahminlerin yaklaşık olarak %96’sı gerçekten pozitiftir. En yüksek recall (duyarlılık) değeri ise “rock” müzik türünde 0.86 ile elde edilmiştir. Pozitif tahmin edilmesi beklenen durumların yaklaşık %86’sı doğru tahmin edilmiştir. En yüksek F1-Skor ise 0.89 ile “rap” müzik türünde elde edilmiştir. Ağırlıklı ortalaması en yüksek olan metrik “kesinlik” olarak görülmüştür.

SVM ile arabesk müzik türü için tüm mesafe fonksiyonlarından elde edilen en yüksek kesinlik, duyarlılık ve F1-skor değerleri aşağıda verilmiştir (Çizelge 5.13).

**Çizelge 5.13** : SVM ile arabesk müzik türü için çekirdek fonksiyonlarının en yüksek başarımları.

Tür : Arabesk	Precision	Recall	F1-Score
Radyal Tabanlı Çekirdek	0.84	0.68	0.75
Polinomal Çekirdek	0.45	0.81	0.58
Lineer Çekirdek	0.69	0.65	0.67

SVM algoritması ile tüm çekirdek fonksiyonları kullanılarak “Arabesk” müzik türünde elde edilen en yüksek değer “kesinlik” için 0.84, “duyarlılık” için 0.81, “F1 Skor” için ise yine 0.75 olarak görülmüştür.

SVM ile dini müzik türü için tüm mesafe fonksiyonlarından elde edilen en yüksek kesinlik, duyarlılık ve F1-skor değerleri aşağıda verilmiştir (Çizelge 5.14).

**Çizelge 5.14** : SVM ile dini müzik türü için çekirdek fonksiyonlarının en yüksek başarımları.

Tür : Dini	Precision	Recall	F1-Score
Radyal Tabanlı Çekirdek	0.54	0.70	0.61
Polinomal Çekirdek	0.61	0.41	0.49
Linear Çekirdek	0.46	0.59	0.65

SVM algoritması ile tüm çekirdek fonksiyonları kullanılarak “Dini” müzik türünde elde edilen en yüksek değer “kesinlik” için 0.61, “duyarlılık” için 0.70, “F1 Skor” için ise yine 0.65 olarak görülmüştür.

SVM ile klasik müzik türü için tüm mesafe fonksiyonlarından elde edilen en yüksek kesinlik, duyarlılık ve F1-skor değerleri aşağıda verilmiştir (Çizelge 5.15).

**Çizelge 5.15** : SVM ile klasik müzik türü için çekirdek fonksiyonlarının en yüksek başarımları.

Tür : Klasik	Precision	Recall	F1-Score
Radyal Tabanlı Çekirdek	0.65	0.62	0.63
Polinomal Çekirdek	0.71	0.53	0.61
Linear Çekirdek	0.67	0.62	0.65

SVM algoritması ile tüm çekirdek fonksiyonları kullanılarak “Klasik” müzik türünde elde edilen en yüksek değer “kesinlik” için 0.71, “duyarlılık” için 0.62, “F1 Skor” için ise yine 0.65 olarak görülmüştür.

SVM ile pop müzik türü için tüm mesafe fonksiyonlarından elde edilen en yüksek kesinlik, duyarlılık ve F1-skor değerleri aşağıda verilmiştir (Çizelge 5.16).

**Çizelge 5.16** : SVM ile pop müzik türü için çekirdek fonksiyonlarının en yüksek başarımları.

Tür : Pop	Precision	Recall	F1-Score
Radyal Tabanlı Çekirdek	0.96	0.79	0.87
Polinomal Çekirdek	0.95	0.55	0.69
Lineer Çekirdek	0.79	0.79	0.79

SVM algoritması ile tüm çekirdek fonksiyonları kullanılarak “Pop” müzik türünde elde edilen en yüksek değer “kesinlik” için 0.96, “duyarlılık” için 0.79, “F1 Skor” için ise yine 0.87 olarak görülmüştür.

SVM ile rap müzik türü için tüm mesafe fonksiyonlarından elde edilen en yüksek kesinlik, duyarlılık ve F1-skor değerleri aşağıda verilmiştir (Çizelge 5.17).

**Çizelge 5.17** : SVM ile rap müzik türü için çekirdek fonksiyonlarının en yüksek başarımları.

Tür : Rap	Precision	Recall	F1-Score
Radyal Tabanlı Çekirdek	0.94	0.94	0.94
Polinomal Çekirdek	1.00	0.58	0.73
Lineer Çekirdek	0.96	0.82	0.89

SVM algoritması ile tüm çekirdek fonksiyonları kullanılarak “Rap” müzik türünde elde edilen en yüksek değer “kesinlik” için 1.00, “duyarlılık” için 0.94, “F1 Skor” için ise yine 0.94 olarak görülmüştür.

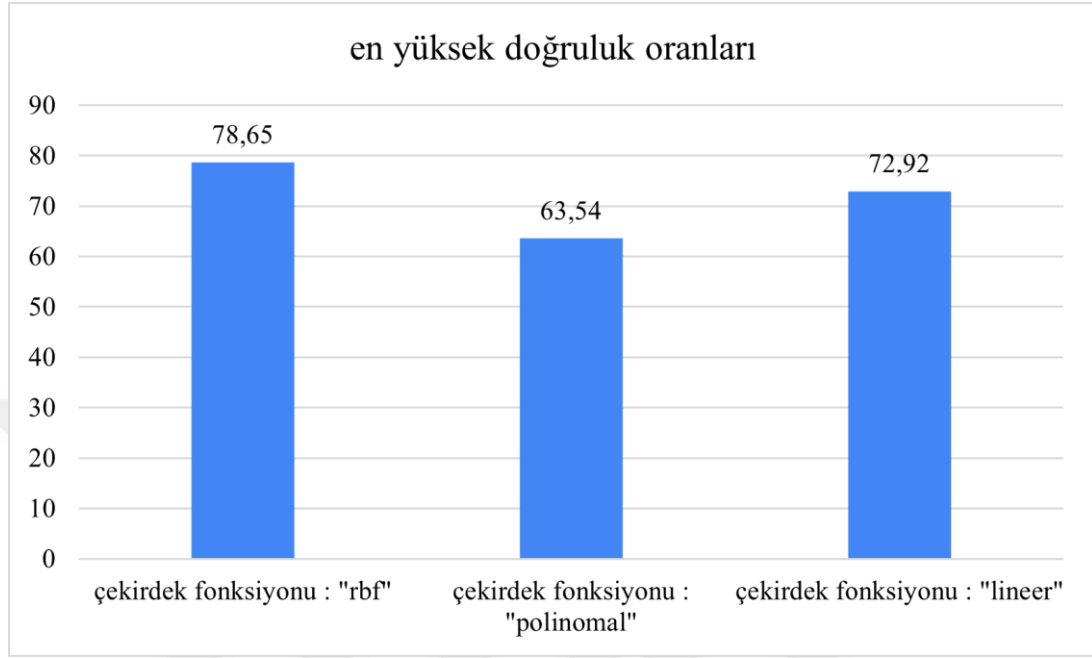
SVM ile rock müzik türü için tüm mesafe fonksiyonlarından elde edilen en yüksek kesinlik, duyarlılık ve F1-skor değerleri aşağıda verilmiştir (Çizelge 5.18).

**Çizelge 5.18** : SVM ile rock müzik türü için çekirdek fonksiyonlarının en yüksek başarımları.

Tür : Rock	Precision	Recall	F1-Score
Radyal Tabanlı Çekirdek	0.83	0.94	0.88
Polinomal Çekirdek	0.83	0.94	0.88
Lineer Çekirdek	0.84	0.86	0.85

SVM algoritması ile tüm çekirdek fonksiyonları kullanılarak “Rock” müzik türünde elde edilen en yüksek değer “kesinlik” için 0.84, “duyarlılık” için 0.94, “F1 Skor” için ise yine 0.88 olarak görülmüştür.

SVM algoritması ile çekirdek fonksiyonu seçimine göre elde edilen en yüksek doğruluk oranları aşağıdaki grafikte verilmiştir (Şekil 5.15).



**Şekil 5.15 :** SVM algoritması ile uygulanan çekirdek fonksiyonları için elde edilen en yüksek doğruluk oranları.

SVM algoritması ile gerçekleştirilen müzik türü sınıflandırmasında her bir çekirdek fonksiyonunda, farklı özniteliklerle en yüksek doğruluk oranları yakalanmıştır. Radyal tabanlı çekirdek fonksiyonunda özniteliklerden tümü sınıflandırma için kullanılarak %78.65 doğruluk oranı, polinomal çekirdek fonksiyonu için "rms" çıkarıldığı durumda %63.54 doğruluk oranı, lineer çekirdek fonksiyonunda ise "spectral\_rolloff" çıkarıldığı durumda %72.92 doğruluk oranı elde edilmiştir.

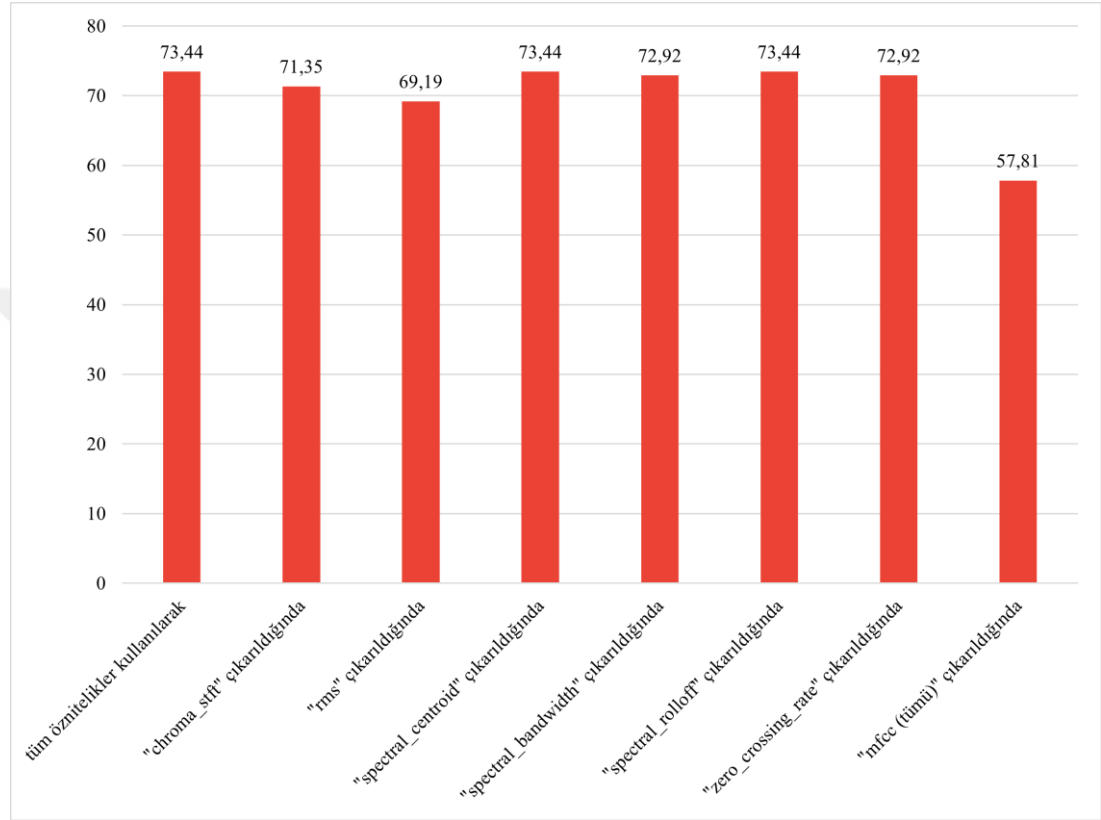
Müzik türlerini sınıflandırmada en yüksek doğruluk oranı, radyal tabanlı çekirdek fonksiyonu kullanılarak elde edilmiştir. Birçok sınıflandırma probleminde popüler durumda olan radyal tabanlı çekirdek fonksiyonunun, SVM ile müzik türü sınıflandırmasında da iyi bir seçenek olduğu görülmüştür.

## 5.6 Lojistik Regresyon Algoritması ile Müzik Türlerinin Sınıflandırılması

Son olarak lojistik regresyon algoritması ile müzik türlerinin sınıflandırılması uygulaması yapılmıştır. Lojistik regresyon oluşturulan modelde "C" hiperparametresine çeşitli denemeler sonucunda en uygun değer olarak tespit edilen

0.1 değeri atanmıştır. Maksimum iterasyon sayısı 100 olarak belirlenmiştir. Bu şekilde oluşturulan model ile elde edilen sonuçlar aşağıda verilmiştir.

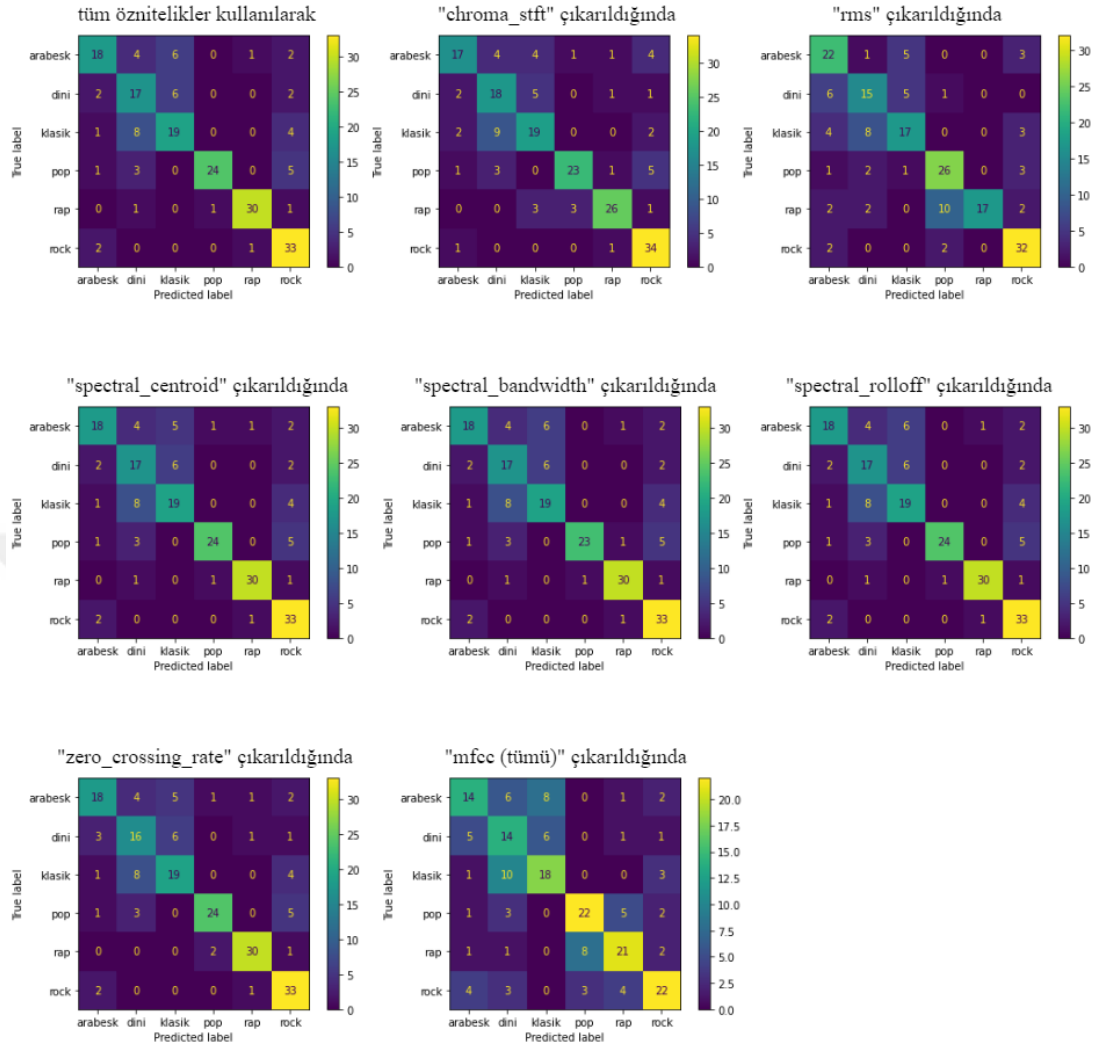
Aşağıdaki şekilde, lojistik regresyon algoritması ile tüm özniteliklerin doğruluk oranlarına etkisinin grafiksel gösterimi verilmiştir. (Şekil 5.16).



**Şekil 5.16 :** Lojistik regresyon algoritması ile elde edilen doğruluk oranları.

Lojistik regresyon algoritması ile uygulanan modelde en yüksek başarımın tüm özniteliklerin kullanıldığı durumda %73.94 doğruluk oranı ile elde edildiği görülmüştür. Ancak "spectral\_centroid" ve "spectral\_rolloff" öznitelikleri ayrı ayrı çıkarıldığı durumda başarıma bir etkileri olmadığı görülmüştür. Diğer tüm öznitelikler dahil edilmediği durumda başarıım düşmüştür. Dolayısıyla kalan özniteliklerin tümünün başarıma pozitif yönde etkisi vardır. Başarımdaki en büyük düşüş oranı, diğer algoritmalarda da benzer şekilde görüldüğü üzere "mfcc" özneliği çıkarıldığı zaman %15.63 gibi bir farklılıkla görülmüştür.

Lojistik regresyon algoritması ile özniteliklerin tüm durumları için tek tek karmaşıklık matrislerinden elde edilen görüntülerin birleştirilmesi ile elde edilen görsel aşağıda verilmiştir (Şekil 5.17).



Şekil 5.17 : Lojistik regresyon algoritması ile elde edilen karmaşıklık matrisleri.

Lojistik regresyon algoritması kullanıldığı durumda elde edilen karmaşıklık matrislerinden bu algoritmanın; “arabesk”, “dini” ve “klasik” müziği sınıflandırmada nispeten başarısız olduğu “rap” ve “rock” müzik türlerini sınıflandırmada ise daha başarılı sonuçlar verdiği görülmektedir.

Lojistik regresyon algoritmasında da SVM algoritmasında elde edilen sonuçlara benzer şekilde en çok “arabesk”, “dini” ve “klasik” müzik türlerinin birbirinden ayırt edilmekte zorlandığı görülmüştür.

Lojistik regresyon algoritması ile en yüksek başarıyı veren öznitelik seçimi ile sınıflandırılan türlerin kesinlik, duyarlılık ve F1-skor değerleri aşağıda verilmiştir (Çizelge 5.19).

**Çizelge 5.19** : Lojistik regresyon algoritması ile elde edilen en yüksek başarıyı veren öznelik seçimi ile sınıflandırılan türlerin başarımlarını gösteren metrikler.

Metrikler	Precision	Recall	F1-Score
arabesk	0.75	0.58	0.65
dini	0.52	0.63	0.57
klasik	0.61	0.59	0.60
pop	0.96	0.73	0.83
rap	0.94	0.91	0.92
rock	0.70	0.92	0.80
Ağırlıklı Ortalama	0.75	0.73	0.74

Lojistik regresyon algoritması ile elde edilen sonuçlarda; en yüksek precision (kesinlik) değeri “pop” müzik türünde 0.94 ile elde edilmiştir. Pozitif tahminlerin yaklaşık olarak %94’ü gerçekten pozitiftir. En yüksek recall (duyarlılık) değeri ise “rock” müzik türünde 0.92 ile elde edilmiştir. Pozitif tahmin edilmesi beklenen durumların yaklaşık %92’si doğru tahmin edilmiştir. En yüksek F1-Skor ise 0.92 ile “rap” müzik türünde elde edilmiştir. Ağırlıklı ortalaması en yüksek olan metrik ise “kesinlik” olarak görülmüştür.

## 6. SONUÇLAR VE ÖNERİLER

Müzik bilgi erişiminde (MIR) ses verileri işlenerek makine öğrenmesi veya derin öğrenme algoritmaları kullanılarak çeşitli yöntemlerle müzikten anlamlı bilgiler çıkarılmakta veya üretilmektedir. Yapılan tez çalışmasında Türkçe müziklerden, ses verilerine ait sinyallerden elde edilen özellikler kullanılarak müzik sınıflandırması gerçekleştirilmiştir.

Sınıflandırmada kullanılan özniteliklerden başarıma en fazla etki sağlayan MFCCs olarak tespit edilmiştir. Bununla birlikte en az etkiyi gösteren her algoritma için farklılık gösterdiği için burada genelleme yapmak mümkün değildir.

Seçilen 6 farklı müzik türlerinin sınıflandırılması sonucu arabesk, klasik ve dini müzik türlerinin birbiriyle yakından ilişkili olduğu gözlemlenmiş olup sınıflandırma algoritmalarının en çok bunları birbirinden ayırt etmede zorlandığı ve hata oranının en yüksek bu türler arasında olduğu görülmüştür. Uygulanan algoritmalarından elde edilen sonuçlarda SVM algoritmasının, seçilen özniteliklere bağlı olmak kaydıyla genel kapsamda en yüksek başarıma sahip olduğu ve bu veri setini sınıflandırmak için en uygun algoritma olduğu tespit edilmiştir. SVM algoritmasındaki en yüksek başarıma radyal tabanlı çekirdek fonksiyonu ile %78.65 doğruluk oranı ile elde edilmiştir. K-NN'de ve lojistik regresyonda nispeten birbirine yakın sonuçlar elde edilmiş olup lojistik regresyon algoritmasının genel başarımasının K-NN'den daha yüksek olduğu görülmüştür. Müzik türü sınıflandırmasında kullanılan özniteliklere göre algoritmaların farklı davranışlar sergilediği tespit edilmiştir. Kullanılan 3 sınıflandırma algoritmasında da özniteliklerin başarıma etkisi farklı oranlarda gözlemlenmiştir. Literatürdeki çalışmalara bakıldığında çeşitli algoritmalar ile MFCCs özellikleri ve ek olarak bazı benzer öznitelikler kullanılarak çeşitli türlere müzik sınıflandırması yapıldığı ve bazı durumlarda daha yüksek başarımlar elde edildiği görülmektedir. Ancak bu tez çalışmasında sınıflandırma için belirlenen tüm algoritmalar ile seçilen 7 farklı özneliğin her birinin ayrı ayrı başarıma etkisi gözlemlenmiştir. Bunun dışında tamamı Türkçe müzik parçalarından oluşan özgün bir veri seti hazırlanıp kullanılmıştır. Bu açıdan literatüre sunulan özgün bir çalışmadır.

Bu tez çalışmasında makine öğrenmesi algoritmaları kullanılarak farklı özniteliklerin müzik türü sınıflandırmasına etkileri gözlemlenmiştir. Yapılan çalışmaya katkı olarak müzik türlerinin birbiriyle olan ilişkisi kullanılarak müzik türü sınıflandırması gerçekleştirilebilir. Örneğin çalışmada kullanılan parçalardan “rock” ve “pop” türlerinin her ikisine de dahil edilebilecek olduğu düşünülen parçalar için “pop-rock” başlığı altında farklı bir tür ilave edilebilir ve daha geniş bir tür yelpazesi elde edilebilir. Bu durumun sınıflandırma başarısına etkisi incelenebilir. Türlerle göre kullanılan enstrümanların ağırlıkları hesaplanıp baskın enstrüman tespiti yapılarak enstrümana göre sınıflandırma çalışması yapılabilir.

Bazı türler kendi içinde de alt türlere ayrılabilir. Seçilen tek bir tür üzerinden veya birden fazla tür üzerinden her bir tür alt kategorilere ayrılarak kendi içinde bir müzik türü sınıflandırması çalışması yapılabilir. Örneğin “rock” müzik için “Anadolu rock”, “alternatif rock” gibi alt türler türetilir.

Birbirine en fazla karışan türleri daha doğru bir şekilde ayırt edebilmek için kullanılan öznitelikler dışında farklı öznitelikler üzerinde çalışılarak başarımları artırılabilir.

Gerçekleştirilen tez çalışmasında yalnızca makine öğrenme algoritmaları kullanılmıştır. Buna ek olarak yapılacak müzik türü sınıflandırması için derin öğrenme algoritmaları kullanılarak veya farklı öznitelikler eklenerek sınıflandırma başarımları artırılabilir.

## KAYNAKLAR

- [1] Xu, C., Maddage, N. C., & Shao, X. (2005). Automatic music classification and summarization. *IEEE transactions on speech and audio processing*, 13(3), 441-450.
- [2] McKinney, M., & Breebaart, J. (2003). Features for audio and music classification.
- [3] Vishnupriya, S., & Meenakshi, K. (2018). Automatic music genre classification using convolution neural network. In *2018 International Conference on Computer Communication and Informatics (ICCCI)* (pp. 1-4). IEEE.
- [4] Durdağ, Z., & Erdoğan, P. (2019). Müzik türlerinin derin öğrenme ağları ile sınıflandırılması. *Sakarya University Journal of Computer and Information Sciences*, 2(1), 53-60.
- [5] Aguiar, R. L., Costa, Y. M., & Silla, C. N. (2018). Exploring data augmentation to improve music genre classification with convnets. In *2018 International Joint Conference on Neural Networks (IJCNN)* (pp. 1-8). IEEE.
- [6] Zhao, W., Zhou, Y., Tie, Y., & Zhao, Y. (2018). Recurrent neural network for MIDI music emotion classification. In *2018 IEEE 3rd Advanced Information Technology, Electronic and Automation Control Conference (IAEAC)* (pp. 2596-2600). IEEE.
- [7] Panwar, S., Das, A., Roopaei, M., & Rad, P. (2017). A deep learning approach for mapping music genres. In *2017 12th System of Systems Engineering Conference (SoSE)* (pp. 1-5). IEEE.
- [8] Guo, J., Liu, A., & Xiao, J. (2020). Melody Classification Based on Performance Event Vector and BRNN. *arXiv preprint arXiv:2010.07562*.
- [9] Jawaharlalnehru, G., & Jothilakshmi, S. (2018). Music Genre Classification using Deep Neural Networks. *International Journal of Scientific Research in Science, Engineering and Technology*, 4(4), 935.
- [10] Kızrak, M. A., & Bolat, B. (2015). Classification of classic Turkish music makams by using deep belief networks. In *2015 23rd Signal Processing and Communications Applications Conference (SIU)* (pp. 527-530). IEEE.
- [11] Ritter, C., Altenhofen, C., Zeppelzauer, M., Kuijper, A., Schreck, T., & Bernard, J. (2018). Personalized Visual-Interactive Music Classification. In *EuroVA@ EuroVis* (pp. 31-35).
- [12] Silla Jr, C. N., Koerich, A. L., & Kaestner, C. A. (2008). Feature selection in automatic music genre classification. In *2008 Tenth IEEE International Symposium on Multimedia* (pp. 39-44). IEEE.

- [13] **Karatana, A., & Yildiz, O.** (2017). Music genre classification with machine learning techniques. In *2017 25th Signal Processing and Communications Applications Conference (SIU)* (pp. 1-4). IEEE.
- [14] **Thiruvengatanadhan, R.** (2018). Music Classification using MFCC and SVM. *International Research Journal of Engineering and Technology*, 5, 922-924.
- [15] **Şimşekli, U.** (2010). Automatic music genre classification using bass lines. In *2010 20th International Conference on Pattern Recognition* (pp. 4137-4140). IEEE.
- [16] **McKay, C., & Fujinaga, I.** (2004). Automatic Genre Classification Using Large High-Level Musical Feature Sets. In *International Society for Music Information Retrieval (ISMIR)* (Vol. 2004, No. 2004, pp. 525-530).
- [17] **Kotsifakos, A., Kotsifakos, E. E., Papapetrou, P., & Athitsos, V.** (2013). Genre classification of symbolic music with SMBGT. In *Proceedings of the 6th international conference on Pervasive technologies related to assistive environments* (pp. 1-7).
- [18] **Richards, T. L., Gates, G. A., Gardner, J. C., Merrill, T., Hayes, C. E., Panagiotides, H., ... & Rubel, E. W.** (1997). Functional MR spectroscopy of the auditory cortex in healthy subjects and patients with sudden hearing loss. *American journal of neuroradiology*, 18(4), 611-620.
- [19] **Alpaydin, E.** (2020). Introduction to machine learning. MIT press. Retrieved from <http://books.google.com/books>
- [20] **Weng, S. F., Reys, J., Kai, J., Garibaldi, J. M., & Qureshi, N.** (2017). Can machine-learning improve cardiovascular risk prediction using routine clinical data?. *PloS one*, 12(4), e0174944.
- [21] **Singh, A., Thakur, N., & Sharma, A.** (2016). A review of supervised machine learning algorithms. In *2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom)* (pp. 1310-1315). IEEE.
- [22] **Hastie, T., Tibshirani, R., & Friedman, J.** (2009). Unsupervised learning. In *The elements of statistical learning* (pp. 485-585). Springer, New York, NY.
- [23] **Reddy, Y. C. A. P., Viswanath, P., & Reddy, B. E.** (2018). Semi-supervised learning: A brief review. *International Journal of Engineering & Technology*, 7(1.8), 81.
- [24] **Bachman, P., Sordoni, A., & Trischler, A.** (2017). Learning algorithms for active learning. In *international conference on machine learning* (pp. 301-310). *Proceedings of Machine Learning Research*.
- [25] **Dietterich, T.** (1995). Overfitting and undercomputing in machine learning. *ACM computing surveys (CSUR)*, 27(3), 326-327.
- [26] **Ying, X.** (2019). An overview of overfitting and its solutions. In *Journal of Physics: Conference Series* (Vol. 1168, No. 2, p. 022022). IOP Publishing.

- [27] **Leondes, C. T.** (1996). *Multidimensional Systems Signal Processing Algorithms and Application Techniques: Advances in Theory and Applications*. Elsevier.
- [28] **Kaski, P., Östergård, P. R., & Patric, R. J.** (2006). *Classification algorithms for codes and designs (Vol. 15)*. Berlin: Springer.
- [29] **Olson, N. D., Lund, S. P., Colman, R. E., Foster, J. T., Sahl, J. W., Schupp, J. M., ... & Zook, J. M.** (2015). Best practices for evaluating single nucleotide variant calling methods for microbial genomics. *Frontiers in genetics*, 6, 235.
- [30] **Hand, D. J.** (2007). Principles of data mining. *Drug safety*, 30(7), 621-622. Springer
- [31] **Wang, H.** (2002). Nearest neighbours without k: A classification formalism based on probability. Faculty of Informatics, University of Ulster.
- [32] **Shivashankara, S., & Srinath, S.** (2019). Shape Based Continuous Real Time Hand Gesture Recognition System of American Sign Language using KNN Classifier.
- [33] **Vapnik, V.** (1999). *The nature of statistical learning theory*. Springer science & business media. Retrieved from <http://books.google.com/books>
- [34] **Oowski, S., Siwek, K., & Markiewicz, T.** (2004). MLP and SVM networks-a comparative study. In *Proceedings of the 6th Nordic Signal Processing Symposium, 2004. NORSIG 2004*. (pp. 37-40). IEEE.
- [35] **Battineni, G., Chintalapudi, N., & Amenta, F.** (2019). Machine learning in medicine: Performance calculation of dementia prediction by support vector machines (SVM). *Informatics in Medicine Unlocked*, 16, 100200.
- [36] **Kavzoglu, T., & Colkesen, I.** (2009). A kernel functions analysis for support vector machines for land cover classification. *International Journal of Applied Earth Observation and Geoinformation*, 11(5), 352-359.
- [37] **Lee, C., & Park, S.** (2011). Damage classification of pipelines under water flow operation using multi-mode actuated sensing technology. *Smart Materials and Structures*, 20(11), 115002.
- [38] **Patle, A., & Chouhan, D. S.** (2013). SVM kernel functions for classification. In *2013 International Conference on Advances in Technology and Engineering (ICATE)* (pp. 1-9). IEEE.
- [39] **Ben-Hur, A., & Weston, J.** (2010). A user's guide to support vector machines. In *data mining techniques for the life sciences* (pp. 223-239). Humana Press.
- [40] **Çokluk, Ö.** (2010). Lojistik Regresyon Analizi: Kavram ve Uygulama. *Educational Sciences: Theory & Practice*, 10(3).
- [41] **Şenel, S., & Alatli, B.** (2014). Lojistik regresyon analizinin kullanıldığı makaleler üzerine bir inceleme. *Journal of Measurement and Evaluation in Education and Psychology*, 5(1), 35-52.
- [42] **Lobo, D., Watters, P., & Wu, X. W.** (2010). A new procedure to help system/network administrators identify multiple rootkit infections. In

2010 Second International Conference on Communication Software and Networks (pp. 124-128). IEEE.

- [43] **Battineni, G., Sagaro, G. G., Nalini, C., Amenta, F., & Tayebati, S. K.** (2019). Comparative machine-learning approach: A follow-up study on type 2 diabetes predictions by cross-validation methods. *Machines*, 7(4), 74.
- [44] **Han, W., Chan, C. F., Choy, C. S., & Pun, K. P.** (2006). An efficient MFCC extraction method in speech recognition. In *2006 IEEE international symposium on circuits and systems* (pp. 4-pp). IEEE.
- [45] **Xu, M., Duan, L. Y., Cai, J., Chia, L. T., Xu, C., & Tian, Q.** (2004). HMM-based audio keyword generation. In *Pacific-Rim Conference on Multimedia* (pp. 566-574). Springer, Berlin, Heidelberg.
- [46] **Sahidullah, M., & Saha, G.** (2012). Design, analysis and experimental evaluation of block based transformation in MFCC computation for speaker recognition. *Speech communication*, 54(4), 543-565.
- [47] **Mohamed, A. R.** (2014). *Deep Neural Network Acoustic Models for ASR* (Doctoral dissertation, University of Toronto).
- [48] **Eskidere, Ö., & Ertaş, F.** (2009). Mel frekanslı keştrüm katsayılarındaki değişimlerin konuşmacı tanıma etkisi. *Uludağ University Journal of The Faculty of Engineering*, 14(2).
- [49] **Kinnunen, T.** (2003). Spectral features for automatic text-independent speaker recognition. Licentiate's thesis.
- [50] **Korkmaz, O. E., & Atasoy, A.** (2015). Emotion recognition from speech signal using mel-frequency cepstral coefficients. In *2015 9th International Conference on Electrical and Electronics Engineering (ELECO)* (pp. 1254-1257). IEEE.
- [51] **Cattani, C., & Rushchitski, J. J.** (2007). *Wavelet and Wave Analysis as applied to Materials with Micro or Nanostructure* (Vol. 74). World Scientific.
- [52] **Heideman, M. T., Johnson, D. H., & Burrus, C. S.** (1985). Gauss and the history of the fast Fourier transform. *Archive for history of exact sciences*, 34(3), 265-277.
- [53] **Van Loan, C.** (1992). *Computational frameworks for the fast Fourier transform*. Society for Industrial and Applied Mathematics.
- [54] **Kalhara, P. G., Jayasinghearachchi, V. D., Dias, A. H. A. T., Ratnayake, V. C., Jayawardena, C., & Kuruwitaarachchi, N.** (2017). TreeSpirit: Illegal logging detection and alerting system using audio identification over an IoT network. In *2017 11th International Conference on Software, Knowledge, Information Management and Applications (SKIMA)* (pp. 1-7). IEEE.
- [55] **Bekele, A. J. A. A.** (2016). Cooley-tukey fft algorithms. *Advanced algorithms*.
- [56] **Molau, S., Pitz, M., Schluter, R., & Ney, H.** (2001). Computing mel-frequency cepstral coefficients on the power spectrum. In *2001 IEEE international conference on acoustics, speech, and signal processing. Proceedings* (cat. No. 01CH37221) (Vol. 1, pp. 73-76). IEEE.

- [57] **Kopparapu, S. K., & Laxminarayana, M.** (2010). Choice of Mel filter bank in computing MFCC of a resampled speech. In 10th International Conference on Information Science, Signal Processing and their Applications (ISSPA 2010) (pp. 121-124). IEEE.
- [58] **Yusnita, M. A., Paulraj, M. P., Yaacob, S., Yusuf, R., & Shahrman, A. B.** (2013). Analysis of accent-sensitive words in multi-resolution mel-frequency cepstral coefficients for classification of accents in Malaysian English. *International Journal of Automotive and Mechanical Engineering*, 7, 1053.
- [59] **Strang, G.** (1999). The discrete cosine transform. *Society for Industrial and Applied Mathematics review*, 41(1), 135-147.
- [60] **Jalil, M., Butt, F. A., & Malik, A.** (2013). Short-time energy, magnitude, zero crossing rate and autocorrelation measurement for discriminating voiced and unvoiced segments of speech signals. In 2013 The international conference on technological advances in electrical, electronics and computer engineering (TAECE) (pp. 208-212). IEEE.
- [61] **Benamirouche, N., Boudraa, B., López-Oller, D., & Pérez-Córdoba, J. L.** (2016). Low complexity forward error correction for CELP-type speech coding over erasure channel transmission. *International Journal of Speech Technology*, 19(4), 717-730.
- [62] **Shete, D. S., Patil, S. B., & Patil, S.** (2014). Zero crossing rate and energy of the speech signal of devanagari script. *IOSR Journal of VLSI and Signal Processing*, 4(1), 1-5.
- [63] **Ramalingam, A., & Krishnan, S.** (2005). Gaussian mixture modeling using short time fourier transform features for audio fingerprinting. In 2005 IEEE International Conference on Multimedia and Expo (pp. 1146-1149). IEEE.
- [64] **Wun, S., Horner, A., & Wu, B.** (2014). Effect of spectral centroid manipulation on discrimination and identification of instrument timbres. *Journal of the Audio Engineering Society*, 62(9), 575-583.
- [65] **Rahman, M. M., & Bhuiyan, M. A. A.** (2011). On segmentation and extraction of features from continuous bangla speech including windowing. *International Journal of Applied Research on Information Technology and Computing*, 2(2), 31-40.
- [66] **Stern, R., & Carter, J. L.** (2003). Polyphase Deformation of the Eastern Central Province, Oklahoma Ouachita Mountains: Evidence in Support of Episodic Orogenesis.
- [67] **D'Amico, A. A., D'Andrea, A. N., & Mengali, U.** (1999). Feedforward joint phase and timing estimation with OQPSK modulation. *IEEE Transactions on vehicular technology*, 48(3), 824-832.
- [68] **Kattel, M., Nepal, A., Shah, A. K., & Shrestha, D.** (2019). Chroma feature extraction. In Conference: Chroma Feature Extraction using Fourier Transform.
- [69] **Korzeniowski, F., & Widmer, G.** (2016). Feature learning for chord recognition: The deep chroma extractor. arXiv preprint arXiv:1612.05065.

- [70] **Rajesh, S., & Nalini, N. J.** (2020). Musical instrument emotion recognition using deep recurrent neural network. *Procedia Computer Science*, 167, 16-25.
- [71] **Sandhya, P., Spoorthy, V., Koolagudi, S. G., & Sobhana, N. V.** (2020, December). Spectral Features for Emotional Speaker Recognition. In *2020 Third International Conference on Advances in Electronics, Computers and Communications (ICAECC)* (pp. 1-6). IEEE.



## **EKLER**

**EK A:** <https://github.com/cuneyt-ozbalci/Turkish-Music-Classification-with-SVM-K-NN-and-Logistic-Regression-Algorithms>

## ÖZGEÇMİŞ

**Ad-Soyad** : Mehmet Cüneyt ÖZBALCI

**Doğum Tarihi ve Doğum Yeri** :

**E-posta** :

### ÖĞRENİM DURUMU:

- **Lisans** : 2018, Erciyes Üniversitesi, Mühendislik Fakültesi, Bilgisayar Mühendisliği

### MESLEKİ DENEYİM:

- Maltepe Üniversitesi (Araştırma Görevlisi), 26.03.2019-13.12.2019
- Bursa Teknik Üniversitesi (Araştırma Görevlisi), 30.12.2020-halen çalışıyor